

On Capacity Regions of Discrete Asynchronous Multiple Access Channels

L. Farkas, T. Kóí *

March 2, 2013

Abstract

A general formalization is given for asynchronous multiple access channels which admits different assumptions on delays. This general framework allows the analysis of so far unexplored models leading to new interesting capacity regions. In particular, a single letter characterization is given for the capacity region in case of 3 senders, 2 synchronous with each other and the third not synchronous with them.

1 Introduction

Ahlsweide [1] and Liao [11] showed that if two senders communicate synchronously over a discrete memoryless multiple access channel (MAC) which is characterized by a stochastic matrix $W(y|x_1, x_2)$, it is possible to communicate with arbitrary small average probability of error if the rate pair is in the following pentagon:

$$\begin{aligned} 0 &\leq R_1 \leq I(X_1 \wedge Y|X_2) \\ 0 &\leq R_2 \leq I(X_2 \wedge Y|X_1) \\ R_1 + R_2 &\leq I(X_1, X_2 \wedge Y) \end{aligned} \tag{1}$$

for some independent input random variables X_1, X_2 , where $P(Y = y|X_1 = x_1, X_2 = x_2) = W(y|x_1, x_2)$. Moreover, the convex hull of the union of these pentagons can also be achieved via time sharing, while no rate pair outside this convex hull is achievable.

The discrete memoryless asynchronous multiple access channel (AMAC) arises when the senders can not synchronize the starting times of their codewords, rather, there is an unknown delay between these starting times. Cover, McEliece and Posner [3] showed that if the delay is bounded by b_n depending on the codeword length n such that $\frac{b_n}{n} \rightarrow 0$ then the convex closure is still achievable by a generalized time sharing method.

Poltyrev [14] and Hui and Humblet [10] addressed models with arbitrary delays known (in [14]) or unknown (in [10]) to the receiver. For such models, the capacity region was shown to be the union of the pentagons above. Verdú [15] studied asynchronous channels with memory. His model slightly differs from common models: the time runs over a torus rather than from $-\infty$ to ∞ . Later, Grant, Rimoldi, Urbanke and Whiting in [7] showed that in the informed receiver case the union can be achieved by rate splitting and successive decoding. For the uninformed receiver case, however, the theory of AMAC contained gaps, which remained untouched for a long time, see Appendix A for details. In essence, these gaps were filled in the book of El Gamal and Kim [6].

This paper is an extended version of the ISIT 2011 contribution [5], originating from the authors' effort to derive the AMAC capacity region without gaps in the proof (the result in [6] was unknown to us at the time, as was, apparently, to the reviewers of [5]). More than doing that, in [5] a general formalization for AMACs was introduced, allowing dependence of the capacity region on the distribution of the delays, typically through the support of that distribution. For a particular (somewhat artificial) choice of the delay distribution, the capacity region was determined which, as shown by an example, could be strictly between the union and its convex closure.

The main technical result of this paper is new compared to [5], it was merely announced in the authors' ISIT 2011 talk. It is a single letter characterization of the capacity region for 3 senders case, where two senders are synchronized with each other and the third is unsynchronized with them.

In section 2 we give the formal description of the AMAC model, where several possible definitions are given, which are analyzed in parallel. In Section 3 a general converse is presented. This converse is used when capacity region of known (in section 4) and previously unknown (in sections 5 and 6) models are derived.

Our achievability proofs rely on the techniques of rate splitting and successive decoding developed by Grant, Rimoldi, Urbanke, Whiting [7] and Rimoldi [12].

*This paper has been submitted to IEEE Transactions on Information Theory. It has been presented in part at ISIT 2011, Saint Petersburg. L. Farkas is with the Department of Analysis, Budapest University of Technology and Economics, e-mail: lfarkas@math.bme.hu. T. Kóí is with the Graduate School of Mathematics and Computer Science, Budapest University of Technology and Economics, e-mail: koitomi@math.bme.hu.

2 Model of coding for the AMAC

In this paper vectors (finite sequences) will be denoted by boldface symbols.

A K -senders asynchronous discrete memoryless multiple-access channel (K -AMAC) is defined in terms of K finite input alphabets $\mathcal{X}_i, i \in \{1, 2, \dots, K\}$, a finite output alphabet \mathcal{Y} , and a stochastic matrix $W : \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_K \rightarrow \mathcal{Y}$ describing the probability distribution of the output given the inputs.

Definition 1. A code-book system of block-length n with rate vector $\mathbf{R} = (R_1, R_2, \dots, R_K)$ for a given K -AMAC W consists of K code-books C_1, C_2, \dots, C_K , where the code-book C_m of the m -th sender has 2^{nR_m} codewords of length n whose symbols are from \mathcal{X}_m .

The system is symbol synchronized but not frame synchronized. The differences between the timing of the receiver and the timings of the senders are represented by a K -tuple of delays as in Definition 3.

The senders have two-way infinite sequences of random messages, and assign codewords to their consecutive messages. The codewords go through the channel. The sequences of the senders' codewords and hence also the output symbol sequence are two-way infinite sequences. Fix the location of the 0-th output symbol. The message of sender $m \in \{1, 2, \dots, K\}$ whose codeword affects the 0-th output is denoted by $M_{m,0}$. This restricts the delays to be in the set $\{0, 1, \dots, n-1\}$. Formally, we use the following definitions:

Definition 2. For each integer $j \in \mathbb{Z}$ and for each $m \in \{1, 2, \dots, K\}$ let $M_{m,j}$ be a uniformly distributed random variable taking values in the set $\{1, 2, \dots, 2^{nR_m}\}$. All these random variables are independent of each other. The two-way infinite sequence $\{M_{m,j}, j \in \mathbb{Z}\}$ represents the message flow sent by the m -th sender. For each integer $j \in \mathbb{Z}$ and for each $m \in \{1, \dots, K\}$ let $\mathbf{X}_{m,j}$ denote the $M_{m,j}$ -th codeword in the code-book of sender m . Let $X_{m,nj+i}$ be the i -th symbol of $\mathbf{X}_{m,j}$ where $i \in \{0, 1, \dots, n-1\}$.

Definition 3. For each $n \in \mathbb{Z}^+$, let

$$\mathbf{D}(n) = (D_1(n), D_2(n), \dots, D_K(n))$$

be a K -tuple of random variables, not necessarily independent of each other but independent of all previously defined random variables, taking values in the set $\{0, 1, \dots, n-1\}$. $D_m(n)$ will represent the delay of sender m relative to the receiver's timing. The joint distribution of delays is known to the senders and the receiver. The realizations of the random variables $D_1(n), D_2(n), \dots, D_K(n)$ are not known to the senders and, depending on the model, may be known or unknown to the receiver. The sequence $\mathbf{D} = \{\mathbf{D}(1), \mathbf{D}(2), \dots, \mathbf{D}(n), \dots\}$ will be called the delay system. With a slight abuse of notation, we also write \mathbf{D} instead of $\mathbf{D}(n)$.

Remark 1. Our definition allows arbitrary distributions for the delays for each blocklength n . Clearly, in practical models these distributions can not be arbitrary, but have to satisfy consistence conditions. We have chosen this general model since we think that any practical model can be described this way.

Example 1

For each $n \in \mathbb{Z}^+$ and for each $m \in \{1, 2, \dots, K\}$ $D_m(n)$ has uniform distribution on $\{0, 1, \dots, n-1\}$ and they are independent. Following [10] it is called the totally asynchronous case in the paper.

Example 2

Let $K = 2$, for each $n \in \mathbb{Z}^+$ let $D_1(n), D_2(n)$ be independent random variables uniformly distributed on the even numbers of $\{0, 1, \dots, n-1\}$. It is called the even delays case in the paper.

Example 3

Let $K = 3$, for each $n \in \mathbb{Z}^+$ let $D_1(n) = D_2(n)$ be a random variable uniformly distributed on $\{0, 1, \dots, n-1\}$ and let $D_3(n)$ be a random variable independent of $D_1(n)$ and uniformly distributed on $\{0, 1, \dots, n-1\}$. It is called the partly asynchronous three senders case in the paper.

For fixed n , the output sequence is defined as follows:

Definition 4. Let Y_{nj+i} be the output random variable of the channel with transition matrix W when the inputs are $X_{1,nj+i+D_1(n)}, X_{2,nj+i+D_2(n)}, \dots, X_{K,nj+i+D_K(n)}$ where $i \in \{0, 1, \dots, n-1\}$.

It is possible to define the decoder in several ways. We will consider two different definitions, which give the strongest version of the converse and direct parts of the coding theorems, respectively.

Definition 5. An informed infinite decoder is defined as a function which assigns to each two way infinite output sequence realization $\{y_l, l \in \mathbb{Z}\}$ and each realization of $\mathbf{D}(n) = (D_1(n), D_2(n), \dots, D_K(n)), n \in \mathbb{Z}^+$, a K -tuple of messages $\{\hat{m}_{m,0}, m \in \{1, 2, \dots, K\}\}$.

Definition 6. An uninformed L -block decoder, $L \in \mathbb{Z}^+$, is defined as a function which assigns to each $(2Ln+1)$ -tuple $\{y_l, l \in \{-Ln, \dots, 0, \dots, Ln\}\}$ of possible output realizations a K -tuple of messages $\{\hat{m}_{m,0}, m \in \{1, 2, \dots, K\}\}$.

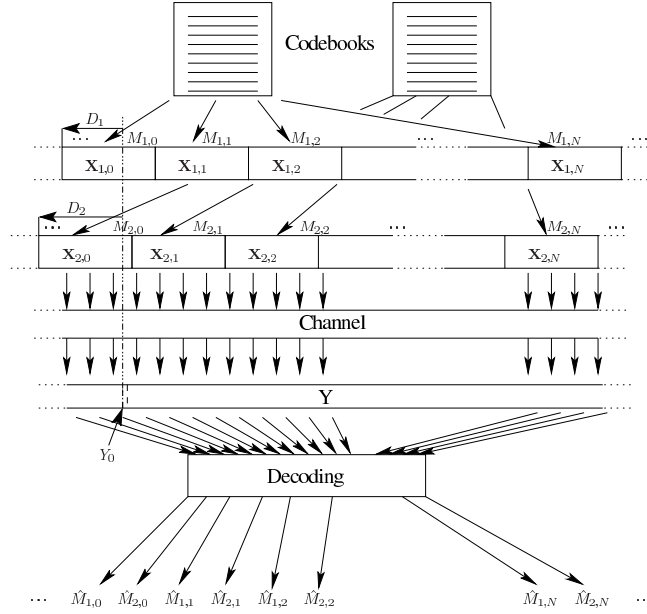


Figure 1: The Setting for Two Senders

The code-books and the decoder form an n -length coding/decoding system.

The definitions above determine the probability structure of the model. For each m the random variable sequence $\{M_{m,j}, j \in \mathbb{Z}\}$ is the two way infinite message flow of the m -th sender. The corresponding flow of codewords is $\{\mathbf{X}_{m,j}, j \in \mathbb{Z}\}$. The flows of the senders, the channel transition and the delay system \mathbf{D} , define a two way infinite output random variable sequence $\{Y_l, l \in \mathbb{Z}\}$. In case of uninformed L -block decoder the receiver examines the output block $Y_{-Ln}, Y_{-Ln+1}, \dots, Y_0, Y_1, \dots, Y_{Ln}$ from which estimations $\{\hat{M}_{m,0}, m \in \{1, 2, \dots, K\}\}$ are created. In case of an informed infinite decoder the whole output sequence and the realizations of delays are used in the estimations $\{\hat{M}_{m,0}, m \in \{1, 2, \dots, K\}\}$. It is assumed that the same but shifted decoding procedure occurs at the output points $\{nk, k \in \mathbb{Z}\}$. Hence the random variables of the estimations $\{\hat{M}_{m,j}, m \in \{1, 2, \dots, K\}, j \in \mathbb{Z}\}$ are also defined. See Fig. 1. for this model, in case $K = 2$.

We will consider two different error definitions. As standard for multiple-access channels, both errors are average error over messages. However, our first error type also involves averaging over delays, while the second one takes maximum over the possible delays.

Definition 7. *The average error is the following:*

$$P_e^n = \Pr \left\{ \bigcup_{m=1}^K \{M_{m,0} \neq \hat{M}_{m,0}\} \right\}. \quad (2)$$

Definition 8. *The maximal error is the following:*

$$P_e^n(*) = \max_{\mathbf{d}(n): \Pr\{\mathbf{D}(n)=\mathbf{d}(n)\} > 0} \Pr \left\{ \bigcup_{m=1}^K \{M_{m,0} \neq \hat{M}_{m,0}\} | \mathbf{D}(n) = \mathbf{d}(n) \right\}. \quad (3)$$

Remark 2. The average error depends on the joint distribution of delays $(D_1(n), D_2(n), \dots, D_K(n))$, while the maximal error depends on the joint distribution of the delays only through its support.

Remark 3. The two kinds of error are related very closely. If $P_e^n(*) \rightarrow 0$ then $P_e^n \rightarrow 0$. On the other hand, if $P_e^n \rightarrow 0$ exponentially as $n \rightarrow \infty$ and if $\min_{\mathbf{d}(n): \Pr\{\mathbf{D}(n)=\mathbf{d}(n)\} > 0} \Pr\{\mathbf{D}(n) = \mathbf{d}(n)\}$ tends to 0 slower than exponentially then also $P_e^n(*) \rightarrow 0$ exponentially.

We have defined several types of models according to the various definitions of decoder and of error. For the sake of brevity, the following definition is meant to define a capacity region simultaneously for all cases.

Definition 9. *Corresponding to the delay system \mathbf{D} , the rate vector (R_1, R_2, \dots, R_K) is achievable if for every $\varepsilon > 0$, $\delta > 0$ for all $N \in \mathbb{Z}^+$ there exists a coding/decoding system with blocklength $n > N$ with rates coordinate-wise exceeding $(R_1 - \delta, R_2 - \delta, \dots, R_K - \delta)$ and with error less than ε . The set of achievable rate vectors is the capacity region of the K -AMAC.*

Remark 4. In the definition above we used the 'optimistic' definition of capacity region, rather than the more usual 'pessimistic one', see [4]¹. The reason is that in the even delays case there are differences in the performance of coding/decoding systems of even and odd blocklength (see Theorem 4).

Remark 5. If for some region achievability is proved in case of uninformed L -block decoder with maximal error, and the converse is proved in case of informed infinite decoder with average error, then for any combination of the model assumptions above the capacity region is equal to this region.

Here, in case of L -block decoder, a proper choice of L is understood. In particular cases, a suitable L will be specified, not entering the question whether a smaller L would also do.

Lemma 1. *For either type of AMAC model, if \mathbf{D} and \mathbf{D}' are two delay systems such that for some $0 < \alpha \leq 1$ for all $n \in \mathbb{Z}^+$ and $\mathbf{d}(n) \in \{0, 1, \dots, n-1\}^K$*

$$\Pr\{\mathbf{D}'(n) = \mathbf{d}(n)\} \geq \alpha \Pr\{\mathbf{D}(n) = \mathbf{d}(n)\}, \quad (4)$$

then the capacity region under delay system \mathbf{D}' is contained (perhaps strictly) in the capacity region under delay system \mathbf{D} .

Proof: Consider an arbitrary n length coding/decoding system. Then $P_{e,\mathbf{D}'(n)}^n \geq \alpha P_{e,\mathbf{D}(n)}^n$ and $P_{e,\mathbf{D}'(n)}^n(*) \geq P_{e,\mathbf{D}(n)}^n(*)$ hold, where the lower indices indicate the underlying delay system. This proves the lemma. ■

Remark 6. In case of any type of decoder, if two delay systems \mathbf{D} and \mathbf{D}' have the same support set for each n , then the capacity regions corresponding to delay systems \mathbf{D} and \mathbf{D}' coincide in case of maximal error. Furthermore, if the equation (4) is fulfilled by \mathbf{D} and \mathbf{D}' and it is also fulfilled when the roles of \mathbf{D} and \mathbf{D}' are reversed, then by Lemma 1 the capacity regions also coincide in case of average error.

3 A general converse

In this section a general converse theorem is proved, which depends on the delay system. In the following sections, this general converse is used to derive the capacity region of special cases.

Write $[i] = \{1, 2, \dots, i\}$. For all subset S of $[K]$ write

$$\mathbf{X}_S = (X_i)_{i \in S} \quad (5)$$

$$S^c = [K] \setminus S, \quad (6)$$

and for all $\mathbf{R} = (R_1, R_2, \dots, R_K)$ write

$$R(S) = \sum_{i \in S} R_i. \quad (7)$$

Let \mathbf{D} denote the delay vector. Let $\mathbf{X}_{B,i+D_B}$ denote the random vector with components $X_{l,i+D_l}$, $l \in B$ where $B \subset [K]$ and $X_{m,j}$ is defined as in definition 2; similar notation is used where $+$ is replaced by \oplus which means addition modulo n .

Theorem 1. *For any n length coding/decoding system for a K senders AMAC W with informed infinite decoder, the following bounds hold for the rate vector $\mathbf{R} = (R_1, R_2, \dots, R_K)$ for all $S \subset [K]$:*

$$R(S) \leq I(\mathbf{X}_{S,Q \oplus D_S} \wedge \tilde{Y}_Q | \mathbf{X}_{S^c,Q \oplus D_{S^c}}, Q, \mathbf{D}) + \varepsilon_n. \quad (8)$$

Here $\varepsilon_n = (R([K]))P_e^n + \frac{1}{n}$, the random variable Q is uniformly distributed on $\{0, 1, \dots, n-1\}$ and independent of \mathbf{D} and the message flows of the senders. Further, \tilde{Y}_Q is linked to the random variables $X_{1,Q \oplus D_1}, X_{2,Q \oplus D_2}, \dots, X_{K,Q \oplus D_K}$ through the channel W ; formally, its conditional distribution given Q, \mathbf{D} and $X_{1,Q \oplus D_1}, X_{2,Q \oplus D_2}, \dots, X_{K,Q \oplus D_K}$ depends only on the values x_1, x_2, \dots, x_K of the latter random variables and is equal to $W(\cdot | x_1, x_2, \dots, x_K)$.

Remark 7. Theorem 1 will be used for sequences of coding/decoding systems with $P_e^n \rightarrow 0$. In this case ε_n also tends to 0.

Proof: For the sake of clarity just the two senders special case is addressed here, the full proof of Theorem 1 can be found in Appendix B. In case of two senders the bounds (8) are:

$$R_1 \leq I(X_{1,Q \oplus D_1} \wedge \tilde{Y}_Q | X_{2,Q \oplus D_2}, Q, D_1, D_2) + \varepsilon_n \quad (9)$$

$$R_2 \leq I(X_{2,Q \oplus D_2} \wedge \tilde{Y}_Q | X_{1,Q \oplus D_1}, Q, D_1, D_2) + \varepsilon_n \quad (10)$$

$$R_1 + R_2 \leq I(X_{1,Q \oplus D_1}, X_{2,Q \oplus D_2} \wedge \tilde{Y}_Q | Q, D_1, D_2) + \varepsilon_n \quad (11)$$

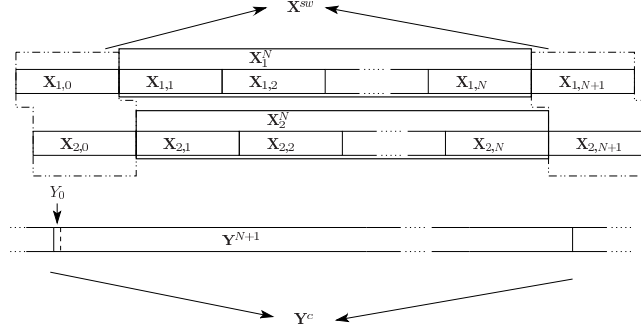


Figure 2: The random variables that play role in the bound on the sum $R_1 + R_2$

Note that $n\varepsilon_n = n(R_1 + R_2)P_e^{(n)} + 1$. Hence

$$n\varepsilon_n \geq H(M_{1,0}, M_{2,0} | \hat{M}_{1,0}, \hat{M}_{2,0}) \quad (12)$$

by Fano's inequality.

We just bound $R_1 + R_2$. The bounds for R_1 and R_2 can be derived similarly (See also Appendix B).

Take a window of the receiver consisting of $N + 1$ n -length blocks $\mathbf{Y}^{N+1} = \{Y_0, Y_1, \dots, Y_{n(N+1)-1}\}$. This window fully covers the code-blocks $\mathbf{X}_{1,1}, \mathbf{X}_{1,2}, \dots, \mathbf{X}_{1,N}$ of sender 1 and $\mathbf{X}_{2,1}, \mathbf{X}_{2,2}, \dots, \mathbf{X}_{2,N}$ of sender 2, denoted by \mathbf{X}_1^N and \mathbf{X}_2^N respectively. The codewords at the sides of the output window are $\mathbf{X}_{1,0}, \mathbf{X}_{2,0}, \mathbf{X}_{1,N+1}, \mathbf{X}_{2,N+1}$, denote this quadruple by \mathbf{X}^{sw} (the expression "side of the windows" is abbreviated by the index sw). Then we have

$$Nn(R_1 + R_2) = \quad (13)$$

$$= H(\mathbf{M}_1^N, \mathbf{M}_2^N) \quad (14)$$

$$= I(\mathbf{M}_1^N, \mathbf{M}_2^N \wedge \hat{\mathbf{M}}_1^N, \hat{\mathbf{M}}_2^N) + \quad (15)$$

$$+ H(\mathbf{M}_1^N, \mathbf{M}_2^N | \hat{\mathbf{M}}_1^N, \hat{\mathbf{M}}_2^N) \leq I(\mathbf{M}_1^N, \mathbf{M}_2^N \wedge \hat{\mathbf{M}}_1^N, \hat{\mathbf{M}}_2^N) + \quad (16)$$

$$+ \sum_{i=1}^N H(M_{1,i}, M_{2,i} | \hat{M}_{1,i}, \hat{M}_{2,i}) \leq I(\mathbf{M}_1^N, \mathbf{M}_2^N \wedge \hat{\mathbf{M}}_1^N, \hat{\mathbf{M}}_2^N) + Nn\varepsilon_n \quad (17)$$

$$\leq I(\mathbf{X}_1^N, \mathbf{X}_2^N \wedge \mathbf{Y}^{N+1}, \mathbf{X}^{sw}, D_1, D_2) + Nn\varepsilon_n \quad (18)$$

where (17) comes from (12) and (18) comes from the Markov relation

$$(\mathbf{M}_1^N, \mathbf{M}_2^N) \ominus (\mathbf{X}_1^N, \mathbf{X}_2^N) \ominus (\mathbf{Y}^{N+1}, \mathbf{X}^{sw}, D_1, D_2) \ominus \quad (19)$$

$\ominus (\mathbf{Y}^{N+1}, \mathbf{Y}^c) \ominus (\hat{\mathbf{M}}_1^N, \hat{\mathbf{M}}_2^N)$. Note that in [5] the Markov relation $(\mathbf{X}_1^N, \mathbf{X}_2^N) \ominus (\mathbf{Y}^{N+1}, D_1, D_2) \ominus (\mathbf{Y}^{N+1}, \mathbf{Y}^c)$ was assumed, which need not hold in general. It seems that Poltyrev [14] also made this error.

Continuing the estimations (13)-(18),

$$Nn(R_1 + R_2) \leq \quad (20)$$

$$\leq I(\mathbf{X}_1^N, \mathbf{X}_2^N \wedge \mathbf{Y}^{N+1}, \mathbf{X}^{sw}, D_1, D_2) + Nn\varepsilon_n = H(\mathbf{X}_1^N, \mathbf{X}_2^N) + Nn\varepsilon_n - \quad (21)$$

$$- H(\mathbf{X}_1^N, \mathbf{X}_2^N | \mathbf{Y}^{N+1}, \mathbf{X}^{sw}, D_1, D_2) = H(\mathbf{X}_1^N, \mathbf{X}_2^N | D_1, D_2) + Nn\varepsilon_n - \quad (22)$$

¹In short, in the 'optimistic' definition it is enough to show that there is a "good" coding/decoding system for a sequence of blocklength $n_k \rightarrow \infty$.

$$\begin{aligned}
&= I(\mathbf{X}_1^N, \mathbf{X}_2^N \wedge \mathbf{Y}^{N+1} | D_1, D_2) + Nn\varepsilon_n + \\
&\quad + I(\mathbf{X}^{sw} \wedge \mathbf{X}_1^N, \mathbf{X}_2^N | \mathbf{Y}^{N+1}, D_1, D_2)
\end{aligned} \tag{23}$$

$$\begin{aligned}
&\leq H(\mathbf{Y}^{N+1} | D_1, D_2) - H(\mathbf{Y}^{N+1} | \mathbf{X}_1^N, \mathbf{X}_2^N, D_1, D_2) + \\
&\quad + 4n \log |\mathcal{X}| + Nn\varepsilon_n
\end{aligned} \tag{24}$$

$$\begin{aligned}
&= H(\mathbf{Y}^{N+1} | D_1, D_2) + 4n \log |\mathcal{X}| + Nn\varepsilon_n - \\
&\quad - \sum_{j=0}^N \sum_{i=0}^{n-1} H(Y_{nj+i} | \mathbf{Y}_0^{nj+i-1} \mathbf{X}_1^N, \mathbf{X}_2^N, D_1, D_2)
\end{aligned} \tag{25}$$

$$\begin{aligned}
&\leq \sum_{j=0}^{(N+1)n-1} H(Y_j | D_1, D_2) + 4n \log |\mathcal{X}| + Nn\varepsilon_n - \\
&\quad - \sum_{j=1}^{N-1} \sum_{i=0}^{n-1} H(Y_{nj+i} | X_{1,nj+i+D_1}, X_{2,nj+i+D_2}, D_1, D_2).
\end{aligned} \tag{26}$$

In (26) we dropped some negative terms (notice that j runs from 1 to $N-1$). Introduce the random variable \tilde{Y}_i linked to the random variables $X_{1,i \oplus D_1}, X_{2,i \oplus D_2}$ by the channel W for all $i \in \{0, 1, \dots, n-1\}$, where \oplus denotes the addition modulo n . For all j the joint distribution of $(D_1, D_2, X_{1,nj+i+D_1}, X_{2,nj+i+D_2}, Y_{nj+i})$ is the same as the joint distribution of $(D_1, D_2, X_{1,i \oplus D_1}, X_{2,i \oplus D_2}, \tilde{Y}_i)$. Using this substitution, (26) can be further bounded from above by:

$$\begin{aligned}
&\leq (N-1) \sum_{i=0}^{n-1} H(\tilde{Y}_i | D_1, D_2) + 2n \log |\mathcal{Y}| + 4n \log |\mathcal{X}| - \\
&\quad - (N-1) \sum_{i=0}^{n-1} H(\tilde{Y}_i | X_{1,i \oplus D_1}, X_{2,i \oplus D_2}, D_1, D_2) + Nn\varepsilon_n
\end{aligned} \tag{27}$$

$$\begin{aligned}
&= (N-1) \sum_{i=0}^{n-1} I(X_{1,i \oplus D_1}, X_{2,i \oplus D_2} \wedge \tilde{Y}_i | D_1, D_2) + \\
&\quad + Nn\varepsilon_n + 4n \log |\mathcal{X}| + 2n \log |\mathcal{Y}|.
\end{aligned} \tag{28}$$

Dividing by nN and introducing the random variable Q uniformly distributed on $\{0, 1, \dots, n-1\}$ and independent of the others we get:

$$\begin{aligned}
R_1 + R_2 &\leq \\
&\leq \frac{N-1}{Nn} \sum_{i=1}^n I(X_{1,i \oplus D_1}, X_{2,i \oplus D_2} \wedge \tilde{Y}_i | D_1, D_2) \\
&\quad + \varepsilon_n + \frac{2 \log |\mathcal{Y}|}{N} + \frac{4 \log |\mathcal{X}|}{N}
\end{aligned} \tag{29}$$

$$\begin{aligned}
&\leq \frac{N-1}{N} I(X_{1,Q \oplus D_1}, X_{2,Q \oplus D_2} \wedge \tilde{Y}_Q | Q, D_1, D_2) \\
&\quad + \varepsilon_n + \frac{2 \log |\mathcal{Y}|}{N} + \frac{4 \log |\mathcal{X}|}{N}
\end{aligned} \tag{30}$$

If $N \rightarrow \infty$ then

$$R_1 + R_2 \leq I(X_{1,Q \oplus D_1}, X_{2,Q \oplus D_2} \wedge \tilde{Y}_Q | Q, D_1, D_2) + \varepsilon_n. \tag{31}$$

■

Corollary 1. *Under the assumptions of Theorem 1 the following bounds hold in the 2-senders case:*

$$R_1 \leq I(X_{1,Q} \wedge \hat{Y}_Q | X_{2,Q \oplus D}, Q, D) + \varepsilon_n \tag{32}$$

$$R_2 \leq I(X_{2,Q \oplus D} \wedge \hat{Y}_Q | X_{1,Q}, Q, D) + \varepsilon_n \tag{33}$$

$$R_1 + R_2 \leq I(X_{1,Q}, X_{2,Q \oplus D} \wedge \hat{Y}_Q | Q, D) + \varepsilon_n. \tag{34}$$

Here Q is uniformly distributed on $\{0, 1, \dots, n-1\}$ and independent of D_1, D_2 and the message flows of the senders, \ominus denotes the subtraction modulo n , $D = D_1 \ominus D_2$ is the relative delay between the two senders and \hat{Y}_Q is linked to $X_{1,Q}, X_{2,Q \oplus D}$ through the channel W .

Proof: Expand the right sides of the equations (9),(10),(11) as sums for the possible values of Q, D_1, D_2 , e.g.

$$\begin{aligned} & \mathbb{I}(X_{1,Q \oplus D_1}, X_{2,Q \oplus D_2} \wedge \tilde{Y}_Q | Q, D_1, D_2) = \\ &= \sum_q \sum_{d_1} \sum_{d_2} \frac{1}{n} \cdot \Pr(D_1 = d_1) \Pr(D_2 = d_2) \mathbb{I}(X_{1,q \oplus d_1}, X_{2,q \oplus d_2} \wedge \tilde{Y}_{q,d_1,d_2}). \end{aligned} \quad (35)$$

Substituting $q' = q \oplus d_1$ and $d = d_1 \oplus d_2$, and renaming q' to q , the Corollary is proved. ■

4 Known capacity regions with a new insight

4.1 The asynchronous one-sender model

In this section the asynchronous model from section 2 is analyzed where there is just one sender ($K = 1$). This will provide the basics for the decoding method of the K-AMAC in general.

In case of $K = 1$, $W : \mathcal{X} \rightarrow \mathcal{Y}$ denotes a classical DMC. For the sake of clarity, we omit from the notations of Section 2 the index corresponding to the unique sender. Let $\{\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(\mathcal{M})\}$ denote the codewords of the code-book of the sender, where $\mathcal{M} = 2^{nR}$ is the number of codewords in the code-book of the sender. The coordinates of $\mathbf{x}(i)$ are denoted by $(x_0(i), x_1(i), \dots, x_{n-1}(i))$.

The difference between this model and the classical one is that the task of the receiver is not just decoding the codewords but also to find the beginning of the codewords.

Theorem 2. *For each version of the model the capacity region of the one sender asynchronous model is the same as that of the classical model: $[0, \max_p(I(p, W))]$, in case of arbitrary delay system.*

Remark 8. By the achievability proof below, beyond decoding the codewords, the receiver also finds out the delay of the sender, with probability exponentially close to 1.

Proof of Theorem 2: The converse part follows from Theorem 1.

In order to prove the direct part it is enough to restrict attention to uninformed L -block decoder and to maximal error; actually $L = 1$ suffices. Classical random argument is used. Let p be an arbitrary distribution over the input alphabet \mathcal{X} . Chose the symbols of codewords in the code-book of rate $0 < R < I(p, W)$ independently according to p . Let $P^n(\mathbf{x}^n, \mathbf{y}^n)$ be the joint distribution on $\mathcal{X}^n \times \mathcal{Y}^n$ induced by the n -th power of p and by the memoryless channel W . Let q^n be the marginal of P^n on \mathcal{Y}^n . We define the decoder as follows. In order to estimate the 0-th sent message $M_{1,0}$, the receiver first examines the n -tuple of outputs $(Y_{-n+1}, Y_{-n+2}, \dots, Y_0)$, then it examines the next n -tuple $(Y_{-n+2}, Y_{-n+3}, \dots, Y_1)$, etc. until the n -tuple $(Y_0, Y_1, \dots, Y_{n-1})$. The estimate will be $\hat{M}_{1,0} = s$ if among the examined n -tuples there is a unique one denoted by Y^n , for which $((X_0(s), X_1(s), \dots, X_{n-1}(s)), Y^n)$ belongs to the typical set

$$S_n^\delta := \left\{ (\mathbf{x}^n, \mathbf{y}^n) : \left| \frac{1}{n} \log \frac{P^n(\mathbf{x}^n, \mathbf{y}^n)}{p^n(\mathbf{x}^n)q^n(\mathbf{y}^n)} - I(p, W) \right| \leq \delta \right\}, \quad (36)$$

and also this s is unique.

It can be assumed that $M_{1,0}$ is fixed, say $M_{1,0} = 1$. It is clear from the classical channel coding theorem that if the decoder examines an n -tuple Y^n which is the output of the whole codeword $\mathbf{X}(1)$, then the decoder will find $\mathbf{X}(1)$ but no other codewords jointly typical with Y^n , with probability exponentially close to 1. Hence we only have to discuss the cases when the decoder examines output symbols in a window of length n , in which one part of the channel input symbols are coming from the 1st codeword and the other part from another codeword r . The probability that $M_{1,-1}$ or $M_{1,1}$ is equal to 1 is exponentially small. Hence it can be assumed that $r \neq 1$. Furthermore, it can be assumed that the input window starts with the r' th codeword as the opposite case is similar. Hence, the channel input symbols in the examined output window can be written as $(X_{n-l}(r), \dots, X_{n-1}(r), X_0(1), \dots, X_{n-l-1}(1))$ for some $r \in \{1, \dots, \mathcal{M}\}, n > l > 0, r \neq 1$. We will show that the probability of incorrectly recognizing typicality in this window is small. The probability, conditioned on the previously presented structure of the examined window, that the s 'th codeword will be typical with this examined output n tuple can be written as:

$$\begin{aligned} & \Pr_{\text{cond}} \{ (X_1(s), \dots, X_n(s), Y_1, \dots, Y_n) \in S_n^\delta \} \\ &= \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta} p^n(\mathbf{x}^n(s)) \cdot \Pr_{\text{cond}} \{ (Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s) \} \end{aligned}$$

$$\begin{aligned}
&= \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta} p^n(\mathbf{x}^n(s)) \frac{q^n(\mathbf{y}^n)}{q^n(\mathbf{y}^n)} \\
&\cdot \Pr_{\text{cond}} \{(Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s)\} \\
&\leq \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta} \frac{2^{-n(I(p,W)-\delta)} P^n(\mathbf{x}^n(s), \mathbf{y}^n)}{q^n(\mathbf{y}^n)} \\
&\cdot \Pr_{\text{cond}} \{(Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s)\} \\
&= 2^{-n(I(p,W)-\delta)} \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta} P^n(\mathbf{x}^n(s) | \mathbf{y}^n) \\
&\cdot \Pr_{\text{cond}} \{(Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s)\}. \tag{37}
\end{aligned}$$

In the above derivation the definition of the set S_n^δ is used, and in the last equation $P^n(\mathbf{x}^n(s) | \mathbf{y}^n)$ denotes the conditional probability induced by the joint distribution P^n . At this point we have to use the structure of the examined window. It is known that in this window the second part of the r -th codeword and the first part of the 1st codeword were sent. We should distinguish three cases: $\{s \neq r, s \neq 1\}$, $\{s \neq r, s = 1\}$, $\{s = r, s \neq 1\}$. The first case follows from the fact that $\Pr_{\text{cond}} \{(Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s)\}$ is equal to $\Pr_{\text{cond}} \{(Y_1, \dots, Y_n) = \mathbf{y}^n\}$. The remaining two cases can be treated very similarly. For the sake of brevity we will demonstrate the case $\{s \neq r, s = 1\}$ when the expression (37) is bounded from above by

$$\begin{aligned}
&\left(2^{-n(I(p,W)-\delta)}\right) \cdot \sum_{\mathbf{x}^n(r) \in \mathcal{X}^n} p^n(\mathbf{x}^n(r)) \sum_{(\mathbf{x}^n(1), \mathbf{y}^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \left[\prod_{h=0}^{n-1} P(x_h(1) | y_h) \right] \\
&\cdot \left[\prod_{h=0}^{l-1} W(y_h | x_{n-l+h}(r)) \right] \left[\prod_{h=l}^{n-1} W(y_h | x_{h-l}(1)) \right]. \tag{38}
\end{aligned}$$

Sum in the following order: $x_{n-1}(1)$, y_{n-1} , $x_{n-2}(1)$, y_{n-2} , \dots , $x_l(1)$, y_l , we get the following expression:

$$\begin{aligned}
&\left(2^{-n(I(p,W)-\delta)}\right) \cdot \sum_{\mathbf{x}^n(r) \in \mathcal{X}^n} p^n(\mathbf{x}^n(r)) \sum_{(\mathbf{x}^l(1), \mathbf{y}^l) \in \mathcal{X}^l \times \mathcal{Y}^l} \left[\prod_{h=0}^{l-1} P(x_h(1) | y_h) \right] \\
&\cdot \left[\prod_{h=0}^{l-1} W(y_h | x_{n-l+r}(r)) \right] = 2^{-n(I(p,W)-\delta)}. \tag{39}
\end{aligned}$$

Indeed, the inner sum is equal to 1, because its terms may be regarded as a joint distribution of a memoryless channel with inputs y_i and outputs $x_i(1)$.

In the above derivation we demonstrated that the probability that the receiver finds the s -th codeword typical with an output n tuple whose input symbols consist of two different codewords, can be bounded from above by $2^{-n(I(p,W)-\delta)}$. Using the union bound over all codewords and over all the n tuples examined by the decoder, gives that the probability of recognizing one of the codewords in a window where the inputs are from two different codewords is less than $n\mathcal{M}2^{-n(I(p,W)-\delta)}$.

The above argument shows that the maximal error over delays of the random code is exponentially small in n . Hence there exists a sequence of deterministic coding-decoding systems with exponentially small maximal error over delays. Optimizing the distribution P , we can see that $\max_p I(p, W)$ is achievable. ■

Remark 9. The crucial step of this proof, the summation in equation (38) is adopted from Gray [8].

Remark 10. Note that using [8] a stronger result can be proved: any sequence of deterministic codes which work well for the classical channel coding model, can be modified to work well for the asynchronous model. Namely, if the same random sync sequence of length $k \approx \log^2(n)$ is appended to each of the original codewords, then with probability tending to 1 it is possible to detect the sync sequence and decode the original codewords. This is how Theorem 2 was proved in the first version of the paper. The authors are indebted to anonymous reviewers for pointing out that the proof becomes somewhat simpler if no sync sequence is used.

4.2 The totally asynchronous case

From this point on, the paper strongly relies on the results of [7] and [12]. Though the reader is assumed familiar with the concepts of successive decoding and rate splitting, the basics will be summarized below.

Let W be a channel with K senders.

Definition 10. The convex polytope $\mathcal{R}[W; p(x_1, x_2, \dots, x_K)]$ is the set of rate tuples $\mathbf{R} \in (\mathbb{R}^+)^K$ such that

$$R(S) \leq I(\mathbf{X}_S \wedge Y | \mathbf{X}_{S^c}), S \subseteq [K], \quad (40)$$

where Y and X_1, X_2, \dots, X_K are connected by the channel W , and the joint distribution of X_1, X_2, \dots, X_K is $p(x_1, x_2, \dots, x_K)$.

Definition 11. Let C denote the following set:

$$C := \bigcup_{p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}} \mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}] \quad (41)$$

where the union is over all product distributions.

Definition 12. The set of rate tuples (R_1, R_2, \dots, R_K) from $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}]$ for which $R([K]) = I(\mathbf{X}_{[K]} \wedge Y)$ is the dominant face of $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}]$. It is denoted by $D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}])$.

Definition 13. We say that (R_1, R_2, \dots, R_K) is dominated by $(\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_K)$ if $R_1 \leq \tilde{R}_1, R_2 \leq \tilde{R}_2, \dots, R_K \leq \tilde{R}_K$.

It can be seen² that the points of $D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}])$ cannot be dominated by other points of $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}]$, but any point from $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}]$ can be dominated by a point from the dominant face.

Remark 11. According to Definition 9, if (R_1, R_2, \dots, R_K) is dominated by an achievable rate vector then (R_1, R_2, \dots, R_K) is also achievable.

Recall from [7] the description of the vertices of $D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}])$. Let $\pi = (\pi_1, \pi_2, \dots, \pi_K)$ be an ordering of $[K]$. For all $i \in [K]$ let $R_{\pi_i}^\pi$ be equal to $I(X_{\pi_i} \wedge Y | \mathbf{X}_{\{\pi_1, \dots, \pi_{i-1}\}})$. For example if $K = 3$, and $\pi = (2, 3, 1)$, then $R_2^\pi = I(X_2 \wedge Y)$, $R_3^\pi = I(X_3 \wedge Y | X_2)$, $R_1^\pi = I(X_1 \wedge Y | X_2, X_3)$. Then the rate vector $\mathbf{R}^\pi = (R_1^\pi, R_2^\pi, \dots, R_K^\pi)$ is a vertex, and all vertices of $D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}])$ can be written in this way with appropriate π . Note that the vertices \mathbf{R}^π need not be all distinct.

In the Appendix of [7] it is proved for informed $L = K$ block decoder that in the totally asynchronous case $\mathbf{R}^\pi \in D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}])$ can be achieved by successive decoding with ordering π . We summarize the proof for $\mathbf{R}^{\{1, 2, \dots, K\}}$. The coding/decoding system is randomly constructed the following way. The symbols of the code-books of the senders are chosen independently according to the appropriate input distributions. The receiver first decodes by joint typicality the codewords of the first sender, considering the random codewords of the other senders as noise. This means that the receiver behaves as if there were only one sender and the channel was the following:

$$W^1(y|x_1) = \sum_{x_2 \in \mathcal{X}_2} \sum_{x_3 \in \mathcal{X}_3} \dots \sum_{x_K \in \mathcal{X}_K} p_{X_2}(x_2)p_{X_3}(x_3) \dots p_{X_K}(x_K)W(y|x_1, x_2, \dots, x_K). \quad (42)$$

Next the receiver decodes the codewords of the second sender by joint typicality using the already decoded codewords of the first sender, considering the other senders as noise. This means that the receiver behaves as it would in a one sender model when the channel was the following:

$$W^2(y, x_1|x_2) = \sum_{x_3 \in \mathcal{X}_3} \sum_{x_4 \in \mathcal{X}_4} \dots \sum_{x_K \in \mathcal{X}_K} p_{X_1}(x_1)p_{X_3}(x_3) \dots p_{X_K}(x_K)W(y|x_1, x_2, \dots, x_K). \quad (43)$$

The codewords of the other senders are decoded similarly. In the final decoding step the receiver decodes the codewords of the K 'th sender by joint typicality using the already decoded codewords of all the other senders. This means that the receiver behaves as it would in a one sender model when the channel was the following:

$$W^K(y, x_1, x_2, \dots, x_{K-1}|x_K) = p_{X_1}(x_1)p_{X_2}(x_2) \dots p_{X_{K-1}}(x_{K-1})W(y|x_1, x_2, \dots, x_K). \quad (44)$$

More detail can be found in the Appendix of [7].

Now recall the notion of individual split from [7]. A split of sender i with input distribution p_{X_i} on $\mathcal{X}_i = \{0, 1, \dots, |\mathcal{X}_i| - 1\}$ results in two virtual senders ia, ib with distributions $p_{X_{ia}}$ and $p_{X_{ib}}$, also on \mathcal{X}_i , explicitly determined by p_{X_i} and a splitting parameter, such that the splitting function $f(x_a, x_b) = \max(x_a, x_b)$ maps $p_{X_{ia}} \times p_{X_{ib}}$ into p_{X_i} .

Section 2 of [7] shows in the totally asynchronous case that each $\mathbf{R} \in D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}])$ can be achieved with Rate Splitting via at most $K - 1$ splits³. This means that a good code for W with rate vector \mathbf{R} can be obtained from a code with successive decoding for an auxiliary channel $W'_\mathbf{R}$ with $2K - 1$ virtual senders constructed by splitting the original senders, perhaps some of them split repeatedly and others not at all; the rate vector of this code equals the vertex \mathbf{R}'^π of the dominant face of $\mathcal{R}[W'_\mathbf{R}; p_{\tilde{X}_1} \times p_{\tilde{X}_2} \times \dots \times p_{\tilde{X}_{2K-1}}]$ for some ordering π and distributions $p_{\tilde{X}_1} \times p_{\tilde{X}_2} \times \dots \times p_{\tilde{X}_{2K-1}}$. In particular, the i 'th coordinate of \mathbf{R} is the sum of those coordinates of \mathbf{R}'^π that correspond to the virtual senders into which the i 'th sender has been split, $i = 1, 2, \dots, K$.

²[7] states it as a consequence of the fact that $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times \dots \times p_{X_K}]$ is a polymatroid, which was observed in [9], [13].

³The stronger result of Section 3 of [7] is not necessary in this paper.

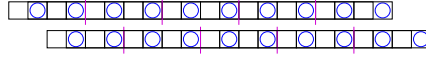


Figure 3: Time sharing when the relative delay is uniform on even numbers

Theorem 3. *In the totally asynchronous case (Example 1), for each model version the capacity region is C .*

Proof: In the converse part it is enough to treat the case of an informed infinite decoder and average error. The right side of eq. (8) can be bounded from above as follows.

$$I(\mathbf{X}_{S, Q \oplus D_S} \wedge \tilde{Y}_Q | \mathbf{X}_{S^c, Q \oplus D_{S^c}}, Q, \mathbf{D}) = \quad (45)$$

$$= H(\tilde{Y}_Q | \mathbf{X}_{S^c, Q \oplus D_{S^c}}, Q, \mathbf{D}) - H(\tilde{Y}_Q | \mathbf{X}_{[K], Q \oplus D_{[K]}}, Q, \mathbf{D}) \quad (46)$$

$$= H(\tilde{Y}_Q | \mathbf{X}_{S^c, Q \oplus D_{S^c}}, Q, \mathbf{D}) - H(\tilde{Y}_Q | \mathbf{X}_{[K], Q \oplus D_{[K]}}) \quad (47)$$

$$\leq H(\tilde{Y}_Q | \mathbf{X}_{S^c, Q \oplus D_{S^c}}) - H(\tilde{Y}_Q | \mathbf{X}_{[K], Q \oplus D_{[K]}}) \quad (48)$$

$$= I(\mathbf{X}_{S, Q \oplus D_S} \wedge \tilde{Y}_Q | \mathbf{X}_{S^c, Q \oplus D_{S^c}}) \quad (49)$$

where (47) comes from the fact that the output depends only on the input variables.

From the fact that the delays are independent and uniform it follows that the random variables $\{Q \oplus D_i, i \in [K]\}$ are independent, hence the random variables $\{X_{i, Q \oplus D_i}, i \in [K]\}$ are also independent. On account of this, the converse statement follows from Theorem 1.

The achievability part needs one modification of the proof in the Appendix of [7] of the assertion that C is achievable with informed $L = 2K - 1$ -block decoder, considering maximal error. In order to get rid of the assumption that the delays are known to the receiver, it is enough to use the synchronization method from Subsection 4.1 in the successive steps of achievability of vertices. ■

Remark 12. In case of two senders Corollary 1 leads to a stronger result. If the relative delay $D = D_1 \ominus D_2$ is uniformly distributed on the set $\{0, 1, \dots, n-1\}$ then, for each model version, the capacity region is C .

5 Even delays

In [6] an artificial but interesting (from theoretical point of view) model is mentioned as open problem: the possible delays are in the set $\{0, 1, \dots, \alpha n\}$ for some $\alpha \in (0, 1)$. In this section, though this problem is not solved, a similar artificial model is analyzed which also has theoretical interest.

Theorem 4. *In the even delays case (Example 2), for each version of the model the capacity region consists of those rate pairs that either belong to C or are linear combinations with weights $\frac{1}{2}, \frac{1}{2}$ of points in C . Moreover, using coding/decoding systems of odd length, only C can be achieved.*

Proof: In order to prove the direct part it is enough to restrict attention to uninformed L -block decoder and to maximal error. Let n be even. Then the senders can do time sharing with weights $\frac{1}{2}, \frac{1}{2}$ using separately the even and the odd symbols and using the coding/decoding method of Theorem 3. Figure 3 demonstrates this fact. Note that in this case L can be chosen as 3.

In the converse part it is enough to treat the case of an informed infinite decoder and average error. The proof uses Corollary 1.

In case of coding/decoding systems of even length the relative delay is uniformly distributed on the even numbers in $\{0, 1, \dots, n-1\}$. Write the upper bounds in Corollary 1 as a sum for the possible values of Q and define two random variables Q_1, Q_2 as uniform on even/odd numbers and independent of each other and everything else. Then the following can be written:

$$R_1 \leq I(X_{1, Q} \wedge \hat{Y}_Q | X_{2, Q \oplus D}, Q, D) + \varepsilon_n \quad (50)$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} I(X_{1, i} \wedge \hat{Y}_i | X_{2, i \oplus D}, D) + \varepsilon_n \quad (51)$$

$$\leq \frac{1}{2} \frac{2}{n} \sum_{i \in \text{odd}} I(X_{1, i} \wedge \hat{Y}_i | X_{2, i \oplus D}) + \frac{1}{2} \frac{2}{n} \sum_{i \in \text{even}} I(X_{1, i} \wedge \hat{Y}_i | X_{2, i \oplus D}) + \varepsilon_n \quad (52)$$

$$\leq \frac{1}{2} I(X_{1, Q_1} \wedge \hat{Y}_{Q_1} | X_{2, Q_1 \oplus D}) + \frac{1}{2} I(X_{1, Q_2} \wedge \hat{Y}_{Q_2} | X_{2, Q_2 \oplus D}) + \varepsilon_n. \quad (53)$$

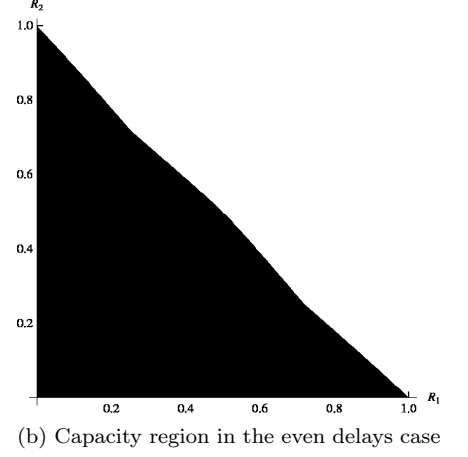
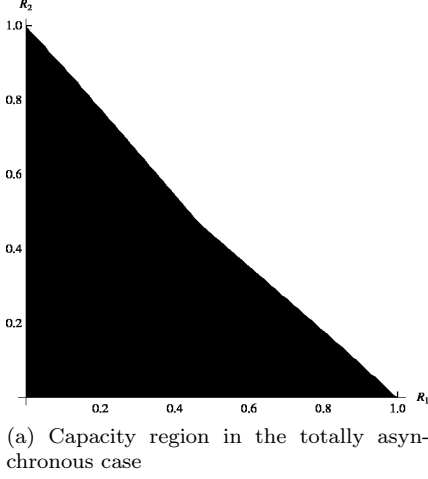


Figure 4: Change of the capacity region with the change in the distribution of the delays

Similarly we get

$$R_2 \leq \frac{1}{2} I(X_{2,Q_1 \oplus D} \wedge \hat{Y}_{Q_1} | X_{1,Q_1}) + \frac{1}{2} I(X_{2,Q_2 \oplus D} \wedge \hat{Y}_{Q_2} | X_{1,Q_2}) + \varepsilon_n \quad (54)$$

$$R_1 + R_2 \leq \frac{1}{2} I(X_{1,Q_1}, X_{2,Q_1 \oplus D} \wedge \hat{Y}_{Q_1}) + \frac{1}{2} I(X_{1,Q_2}, X_{2,Q_2 \oplus D} \wedge \hat{Y}_{Q_2}) + \varepsilon_n \quad (55)$$

where $X_{1,Q_1}, X_{2,Q_1 \oplus D}$ and $X_{1,Q_2}, X_{2,Q_2 \oplus D}$ are independent. This proves the converse result for even blocklength (see [4] Lemma 14.4+, or its generalization, Lemma 2 in Section 6 of this paper).

In the subsequent part of this proof the symbol n denotes odd integer. Now we prove that with coding/decoding systems of odd length, just the union of the pentagons can be achieved. Given such sequence of coding/decoding systems, where $P_e^n \rightarrow 0$, let c_n be a sequence with $c_n \rightarrow 0$ and $\frac{P_e^n}{c_n} \rightarrow 0$ such that $c_n n$ is integer.

Recall that the delays $D_1(n)$ and $D_2(n)$ are independent and uniformly distributed random variables on the set $\{0, 2, \dots, n-1\}$. For all $i \in \{0, 1, \dots, n-1\}$ let $K(i)$ be the number of those pairs $d_1, d_2 \in \{0, 2, \dots, n-1\}$ for which the relative delay $d = d_1 \oplus d_2$ is equal to i , then:

$$K(i) = \begin{cases} \frac{n-i+1}{2} & \text{if } i \text{ is even} \\ \frac{i+1}{2} & \text{if } i \text{ is odd.} \end{cases}$$

Let $D'_1(n)$ and $D'_2(n)$ be two random variables taking values in the set $\{0, 2, \dots, n-1\}$ with the following joint distribution. For each $d_1, d_2 \in \{0, 2, \dots, n-1\}$, if $d_1 \oplus d_2 \in \{c_n n, c_n n + 1, \dots, n-1 - c_n n\}$ let $\Pr\{D'_1(n) = d_1, D'_2(n) = d_2\}$ be equal to $\frac{1}{n(1-2c_n)K(d_1 \oplus d_2)}$, otherwise 0. Then for each $d_1, d_2 \in \{0, 2, \dots, n-1\}$ the following bound holds if n is large enough:

$$\begin{aligned} \frac{4}{(n+1)^2} &= \Pr\{D_1(n) = d_1, D_2(n) = d_2\} \geq \\ &\geq c_n \Pr\{D'_1(n) = d_1, D'_2(n) = d_2\}. \end{aligned} \quad (56)$$

Using the same idea as in the proof of Lemma 1 and the fact that $\frac{P_e^n}{c_n} \rightarrow 0$ we can conclude that the given sequence of coding/decoding systems has average error also tending to 0 under the delay system \mathbf{D}' described by the random variables $D'_1(n)$ and $D'_2(n)$. Hence if we show that under delay system \mathbf{D}' only C can be achieved, the assertion is proved.

Under delay system \mathbf{D}' the relative delay $D'(n) = D'_1(n) \oplus D'_2(n)$ is uniformly distributed on the set $\{c_n n, c_n n + 1, \dots, n-1 - c_n n\}$. By Corollary 1 the following bounds hold for the rates:

$$\begin{aligned} R_1 &\leq I(X_{1,Q} \wedge \hat{Y}_Q | X_{2,Q \oplus D'}, Q, D') + \varepsilon_n \\ R_2 &\leq I(X_{2,Q \oplus D'} \wedge \hat{Y}_Q | X_{1,Q}, Q, D') + \varepsilon_n \\ R_1 + R_2 &\leq I(X_{1,Q}, X_{2,Q \oplus D'} \wedge \hat{Y}_Q | Q, D') + \varepsilon_n. \end{aligned} \quad (57)$$

Let $\bar{D}(n)$ be a random variable uniformly distributed on the set $\{0, 1, \dots, n-1\}$. As the variation distance between the product joint distributions of (Q, D') and (Q, \bar{D}) tends to 0, the following differences also tend to 0 as $n \rightarrow \infty$:

$$\begin{aligned} & I(X_{1,Q}, X_{2,Q \oplus D'} \wedge \hat{Y}_Q | Q, D') - I(X_{1,Q}, X_{2,Q \oplus \bar{D}} \wedge \hat{Y}_Q | Q, \bar{D}), \\ & I(X_{1,Q} \wedge \hat{Y}_Q | X_{2,Q \oplus D'}, Q, D') - I(X_{1,Q} \wedge \hat{Y}_Q | X_{2,Q \oplus \bar{D}}, Q, \bar{D}), \\ & I(X_{2,Q \oplus D'} \wedge \hat{Y}_Q | X_{1,Q}, Q, D') - I(X_{2,Q \oplus \bar{D}} \wedge \hat{Y}_Q | X_{1,Q}, Q, \bar{D}). \end{aligned} \quad (58)$$

Taking into account that $X_{1,Q}$ and $X_{2,Q \oplus \bar{D}}$ are independent, the assertion is proved (See also Remark 12). ■

Example 4

There are two well-known examples ([2], [4]) which show that the convex hull operation can be useful. Here we use [4]. Let the channel be defined by $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}$, $W(0|0, 0) = 1$, $W(1|1, 0) = W(1|0, 1) = 1$ and $W(1|1, 1) = W(0|1, 1) = \frac{1}{2}$. The capacity regions in the totally asynchronous and in the even delays case are shown on Figure 4a and on Figure 4b. In the latter case a hill appears in the middle of the picture.

Remark 13. Remarkably, it does not seem possible to extend this result for uniform relative delay distributions on $\{0, 1, \dots, n/2\}$, although this distribution has the same entropy as the uniform distribution on even (or odd) numbers. Similar results can be achieved if the distribution is uniform on numbers which are divisible by 3. In this case time sharing with weights $\frac{1}{3}, \frac{2}{3}$ becomes possible.

Remark 14. This also means that if the senders of the totally asynchronous AMAC want to time share with weights $(\frac{1}{2}, \frac{1}{2})$, they can do that if a one-shot 1-bit side-information about the delays is available to the senders.

6 Partly asynchronous three-senders case

In this section we will prove coding theorem in case of $K = 3$, when $D_1(n) = D_2(n)$ and $D_3(n)$ are independent and uniformly distributed on the set $\{0, 1, \dots, n-1\}$.

Theorem 5. *In the partially asynchronous three senders case (Example 3), for each version of the model the capacity region is*

$$\bigcup_{p_{X_3}} \text{Conv} \left(\bigcup_{p_{X_1} \times p_{X_2}} \mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}] \right). \quad (59)$$

In words, it consists of the convex combination of rate triples from \mathcal{C} whose corresponding convex polytopes are defined by the same third distribution.

Remark 15. Using the Carathéodory-Fenchel Theorem (e.g. Chapter 15 of [4]) in Theorem 5, it suffices to take convex combinations involving at most three rate triples.

Proof of the converse part in Theorem 5: It is enough to treat the case of an informed infinite decoder and average error. Theorem 1 can be used as follows.

If $S \subset \{1, 2, 3\}$ then the following bound holds:

$$R(S) \leq \sum_{i=1}^n \frac{1}{n} I(\mathbf{X}_{S, i \oplus D_S} \wedge Y_i | \mathbf{X}_{S^c, i \oplus D_{S^c}}, \mathbf{D}) + \varepsilon_n. \quad (60)$$

Summing over the possible values of $D_1 = D_2$ we get the following bounds:

$$\begin{aligned} R_1 & \leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{1, i \oplus d} \wedge \tilde{Y}_i | X_{2, i \oplus d}, X_{3, i \oplus D_3}, D_3) + \varepsilon_n \\ R_2 & \leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{2, i \oplus d} \wedge \tilde{Y}_i | X_{1, i \oplus d}, X_{3, i \oplus D_3}, D_3) + \varepsilon_n \\ R_3 & \leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{3, i \oplus D_3} \wedge \tilde{Y}_i | X_{1, i \oplus d}, X_{2, i \oplus d}, D_3) + \varepsilon_n \end{aligned}$$

$$\begin{aligned}
R_1 + R_2 &\leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{1,i \oplus d}, X_{2,i \oplus d} \wedge \tilde{Y}_i | X_{3,i \oplus D_3}, D_3) + \varepsilon_n \\
R_2 + R_3 &\leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{2,i \oplus d}, X_{3,i \oplus D_3} \wedge \tilde{Y}_i | X_{1,i \oplus d}, D_3) + \varepsilon_n \\
R_1 + R_3 &\leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{1,i \oplus d}, X_{3,i \oplus D_3} \wedge \tilde{Y}_i | X_{2,i \oplus d}, D_3) + \varepsilon_n \\
R_1 + R_2 + R_3 &\leq \sum_{i=0}^{n-1} \sum_{d=0}^{n-1} \frac{1}{n} \frac{1}{n} \cdot I(X_{1,i \oplus d}, X_{2,i \oplus d}, X_{3,i \oplus D_3} \wedge \tilde{Y}_i | D_3) + \varepsilon_n.
\end{aligned} \tag{61}$$

Note that, $X_{3,i \oplus D_3}$ is independent of the other variables, and has the same distribution for all i . Note also that the above inequalities can be overestimated by dropping D_3 from the condition (same argument as in Theorem 3). Hence the converse part follows from Lemma 2 below. ■

The achievability part in Theorem 5 is proved later in this section.

Lemma 2. *Given n sets $\mathcal{R} \left[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i} \right]$, $i \in \{1, 2, \dots, n\}$, a vector (R_1, R_2, R_3) equals a convex combination with weights α_i of n vectors from these sets if and only if they are contained in $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_n)$ which is defined by the following inequalities:*

$$\begin{aligned}
0 &\leq R_1 \leq \sum_{i=1}^n \alpha_i I(X_1^i \wedge Y^i | X_2^i, X_3^i) \\
0 &\leq R_2 \leq \sum_{i=1}^n \alpha_i I(X_2^i \wedge Y^i | X_1^i, X_3^i) \\
0 &\leq R_3 \leq \sum_{i=1}^n \alpha_i I(X_3^i \wedge Y^i | X_1^i, X_2^i) \\
R_1 + R_2 &\leq \sum_{i=1}^n \alpha_i I(X_1^i, X_2^i \wedge Y^i | X_3^i) \\
R_1 + R_3 &\leq \sum_{i=1}^n \alpha_i I(X_1^i, X_3^i \wedge Y^i | X_2^i) \\
R_2 + R_3 &\leq \sum_{i=1}^n \alpha_i I(X_2^i, X_3^i \wedge Y^i | X_1^i) \\
R_1 + R_2 + R_3 &\leq \sum_{i=1}^n \alpha_i I(X_1^i, X_2^i, X_3^i \wedge Y^i).
\end{aligned} \tag{62}$$

Proof: This proof follows the proof of Lemma 14.4+ in [4]. The sets $\mathcal{R} \left[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i} \right]$ and the set $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_n)$ satisfying the above conditions are convex polytopes with 16 vertices. Using the fact that the mutual and the conditional mutual information are always non-negative, it can be easily derived that there are no redundant inequalities between the defining equations of the sets $\mathcal{R} \left[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i} \right]$ and $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_n)$. This means for example that it is not possible that the sum of the bounds for $R_1 + R_2$ and R_3 is strictly less than the bound for $R_1 + R_2 + R_3$. Using this fact the vertices of $\mathcal{R} \left[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i} \right]$ can be written the following way. First $\mathbf{v}_i^0 = (0, 0, 0)$ is a vertex. The remaining 15 vertices can be divided into three groups of equal size. The first group consists of those vertices (R_1, R_2, R_3) for which R_1 is equal to its own bound, i.e., of the vertices:

$$\begin{aligned}
\mathbf{v}_i^1 &= (I(X_1^i \wedge Y^i | X_2^i, X_3^i), 0, 0) \\
\mathbf{v}_i^2 &= (I(X_1^i \wedge Y^i | X_2^i, X_3^i), I(X_2^i \wedge Y^i | X_3^i), 0) \\
\mathbf{v}_i^3 &= (I(X_1^i \wedge Y^i | X_2^i, X_3^i), I(X_2^i \wedge Y^i | X_3^i), I(X_3^i \wedge Y^i)) \\
\mathbf{v}_i^4 &= (I(X_1^i \wedge Y^i | X_2^i, X_3^i), 0, I(X_3^i \wedge Y^i | X_2^i)) \\
\mathbf{v}_i^5 &= (I(X_1^i \wedge Y^i | X_2^i, X_3^i), I(X_2^i \wedge Y^i), I(X_3^i \wedge Y^i | X_2^i))
\end{aligned} \tag{63}$$

The two other groups $(\mathbf{v}_i^6, \mathbf{v}_i^7, \dots, \mathbf{v}_i^{10})$ and $(\mathbf{v}_i^{11}, \mathbf{v}_i^{12}, \dots, \mathbf{v}_i^{15})$ are obtained similarly.

Note that $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$ can be degenerate in the sense that these sixteen vertices need not be all distinct. The vertices of $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_n)$ are the points $\sum_{i=1}^n \alpha_i \mathbf{v}_i^j$, $0 \leq j \leq 15$. As these vertices are contained in the (convex) set of convex combinations with weights α_i of vectors in the sets $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$, therefore whole $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_n)$ is contained. The reverse inclusion is obvious. ■

By Definition 12, the points (R_1, R_2, R_3) of $D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}])$ satisfy the inequalities in (40), with $R_1 + R_2 + R_3 = I(X_1, X_2, X_3 \wedge Y)$. An edge of this dominant face is characterized by another inequality in (40) fulfilled with equality. The set S corresponding to that inequality will be called the type of this edge.

The following lemma states that rate triples lying on edges with a fixed type behave similarly in context of rate splitting and successive decoding. It can be considered as a remark to the general theory of [7], [12] in the special case $K = 3$.

Lemma 3. *For every fixed nonempty $S \subsetneq \{1, 2, 3\}$ there exists a 4-senders channel W' derived from W by splitting the first or the second sender, and an ordering $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$ of the 4 senders with the following property. If W' is derived from W by splitting the first sender, then to any input distributions $p_{X_1}, p_{X_2}, p_{X_3}$ of W and for all $(R_1, R_2, R_3) \in D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}])$ lying on the edge of type S , there exist input distributions $p_{X_{1a}}, p_{X_{1b}}$ and non-negative numbers R_{1a}, R_{1b} with $R_{1a} + R_{1b} = R_1$ such that $(R_{1a}, R_{1b}, R_2, R_3)$ is the vertex of $D(\mathcal{R}[W'; p_{X_{1a}} \times p_{X_{1b}} \times p_{X_2} \times p_{X_3}])$ described by ordering π . If W' is derived from W by splitting the second sender, then to any input distributions $p_{X_1}, p_{X_2}, p_{X_3}$ of W and for all $(R_1, R_2, R_3) \in D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}])$ lying on the edge of type S , there exist input distributions $p_{X_{2a}}, p_{X_{2b}}$ and non-negative numbers R_{2a}, R_{2b} with $R_{2a} + R_{2b} = R_2$ such that $(R_1, R_{2a}, R_{2b}, R_3)$ is the vertex of $D(\mathcal{R}[W'; p_{X_1} \times p_{X_{2a}} \times p_{X_{2b}} \times p_{X_3}])$ described by ordering π .*

Proof: Assume for example that a rate triple $(R_1, R_2, R_3) \in D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}])$ lies on the edge of type $S = \{1, 3\}$, hence $R_1 + R_2 + R_3 = I(X_1, X_2, X_3 \wedge Y)$, $R_1 + R_3 = I(X_1, X_3 \wedge Y|X_2)$. The other cases are similar. Then $R_2 = I(X_2 \wedge Y)$ and (R_1, R_3) lies on $D(\mathcal{R}[\hat{W}; p_{X_1} \times p_{X_3}])$, where $\hat{W}(y, x_2|x_1, x_3) = p_{X_2}(x_2)W(y|x_1, x_2, x_3)$ (see [12], beginning of section 3c). Denote by \hat{W}' the three senders channel derived from \hat{W} by splitting the first sender. We could split the third sender instead of the first sender, but we want to leave the third sender unsplit. Using the basic rate splitting result of [7] for two senders channels, there exist input distributions $p_{X_{1a}}, p_{X_{1b}}$ and non-negative numbers R_{1a}, R_{1b} with $R_{1a} + R_{1b} = R_1$ such that (R_{1a}, R_{1b}, R_3) is the vertex of $D(\mathcal{R}[\hat{W}'; p_{X_{1a}} \times p_{X_{1b}} \times p_{X_3}])$ described by the ordering $(1a, 3, 1b)$. Hence, $(R_{1a}, R_{1b}, R_2, R_3)$ is the vertex of $D(\mathcal{R}[W'; p_{X_{1a}} \times p_{X_{1b}} \times p_{X_2} \times p_{X_3}])$ described by the ordering $(2, 1a, 3, 1b)$, where W' is the 4-senders channel derived from W by splitting the first sender. This argument shows that for $S = \{1, 3\}$, the channel W' derived from W by splitting the first sender, and the ordering $(2, 1a, 3, 1b)$ on the senders of W' fulfill the requirements of this lemma. ■

The next lemma shows that each \mathbf{x} which is not in C but can be written as the convex combination of rate triples from C , can be dominated by a convex combination of rate triples from C which lie on edges of same type.

Lemma 4. *Given k sets $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$, $i \in \{1, 2, \dots, k\}$, if a vector \mathbf{x} is not in C , but can be written as $\mathbf{x} = \sum_{i=1}^k \alpha_i \mathbf{x}_i$, where $\mathbf{x}_i \in \mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$, $0 \leq \alpha_i < 1$, $i \in \{1, 2, \dots, k\}$, $\sum_{i=1}^k \alpha_i = 1$, then \mathbf{x} can be dominated by an \mathbf{x}' which can be written as $\sum_{i=1}^k \alpha'_i \mathbf{x}'_i$, where $\mathbf{x}'_i \in D(\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}])$, $0 \leq \alpha'_i < 1$, $i \in \{1, 2, \dots, k\}$, $\sum_{i=1}^k \alpha'_i = 1$ and the vectors \mathbf{x}'_i , $i \in \{1, 2, \dots, k\}$, lie on edges of same type.*

Proof: It can be assumed that $\alpha_i > 0$ for all i . If \mathbf{x}_i is not on $D(\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}])$ then we can take a dominating $\tilde{\mathbf{x}}_i$ from the dominant face, for all i . Then $\tilde{\mathbf{x}} = \sum_{i=1}^k \alpha_i \tilde{\mathbf{x}}_i$ dominates \mathbf{x} . So it can be assumed that the rate triple \mathbf{x}_i is on $D(\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}])$ for all i .

The dominant face of a set $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$ is a hexagon⁴ on a plane with normal vector $(1, 1, 1)$. We say that the height of the plane with normal vector $(1, 1, 1)$ is a if its equation is $x + y + z = a$. The height of a dominant face is the height of its plane. As in Lemma 2 let us consider the set $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_k)$. This is the set of convex combinations with weights α_i , $1 \leq i \leq k$ of the sets $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$, $1 \leq i \leq k$. The dominant face $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$ of $\mathcal{R}(\alpha_1, \alpha_2, \dots, \alpha_k)$ consists of those points (R_1, R_2, R_3) for which $R_1 + R_2 + R_3 = \sum_{i=1}^k \alpha_i I(X_1^i, X_2^i, X_3^i \wedge Y^i)$. Note that $\mathbf{x} \in \mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$ because the points \mathbf{x}_i are on the dominant face of $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$ respectively. Any edge of $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$ consists of those points for which one of the inequalities (62) is fulfilled with equality. Hence the edges of $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$ consist of points which are convex combinations with weights $\alpha_1, \alpha_2, \dots, \alpha_k$ of points lying on edges of same type. If \mathbf{x} is on an edge of $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$ then we proved the assertion. Hence it can be assumed that \mathbf{x} is an inner point of $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$. Suppose first that there exists m, l such that

⁴The hexagon can be degenerated since some vertices can be identical

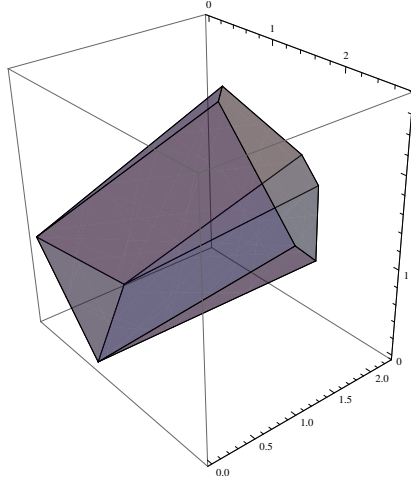


Figure 5: The set of convex combination of two dominant faces. One of them is degenerate (triangle).

$I(X_1^m, X_2^m, X_3^m \wedge Y^m) > I(X_1^l, X_2^l, X_3^l \wedge Y^l)$. Let us define new weights: If $i \neq m, i \neq l$ then let $\alpha'_i = \alpha_i$, and let $\alpha'_m = \alpha_m + \varepsilon$, $\alpha'_l = \alpha_l - \varepsilon$. Then the height of $\mathcal{D}(\alpha'_1, \alpha'_2, \dots, \alpha'_k)$ is larger than the height of $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$. If ε is small then one of the points of $\mathcal{D}(\alpha'_1, \alpha'_2, \dots, \alpha'_k)$ will dominate \mathbf{x} . We increase ε until this property holds or until α'_l becomes 0. Then, using continuity, an edge point of $\mathcal{D}(\alpha'_1, \alpha'_2, \dots, \alpha'_k)$ will dominate \mathbf{x} or $\alpha'_l = 0$ holds. This argument shows that it is enough to restrict attention to the case when $I(X_1^m, X_2^m, X_3^m \wedge Y^m) = I(X_1^l, X_2^l, X_3^l \wedge Y^l)$ for all m, l . This means that the dominant faces of sets $\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}]$ are in the same plane. Using again the continuous change of $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$: if $\alpha_i \rightarrow 1$, and $\alpha_j \rightarrow 0$ if $j \neq i$, then $\mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_k)$ tends to $D(\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}])$. As \mathbf{x} is not in C , it is not in $D(\mathcal{R}[W; p_{X_1^i} \times p_{X_2^i} \times p_{X_3^i}])$ for either $i \in \{1, 2, \dots, k\}$, hence there are weights $\alpha_1^*, \alpha_2^*, \dots, \alpha_k^*$ for which \mathbf{x} is on an edge of $\mathcal{D}(\alpha_1^*, \alpha_2^*, \dots, \alpha_k^*)$. So it is a convex combination of points lying on edges of same type. ■

The proof of achievability in Theorem 5: In order to prove the direct part it is enough to restrict attention to uninformed L -block decoder and to maximal error.

Theorem 3 shows that the rate triples of C can be achieved in the totally asynchronous case considering uninformed $L = 5$ -block decoder with maximal error. It follows that in the partially asynchronous three senders case, C is also achievable considering uninformed $L = 5$ -block decoder with maximal error with the same coding/decoding method. By Remark 15 it is enough to consider points which are not in C but can be written as the convex combination of two or three rate triples from C whose corresponding sets $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}]$ have the same third distribution. Note that the following part of this proof shows that in the latter cases $L = 4$ -block decoder suffices. For the sake of clarity we deal only with points which can be written as the convex combination of two rate triples whose corresponding sets $\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}]$ have the same third distribution. The case of convex combination of three rate triples can be derived similarly.

Let (R_1, R_2, R_3) be in $D(\mathcal{R}[W; p_{X_1} \times p_{X_2} \times p_{X_3}])$ and $(\tilde{R}_1, \tilde{R}_2, \tilde{R}_3)$ in $D(\mathcal{R}[W; p_{\tilde{X}_1} \times p_{\tilde{X}_2} \times p_{X_3}])$. Note that the third input distribution is the same in case of both convex polytopes. We want to show that $\alpha(R_1, R_2, R_3) + (1 - \alpha)(\tilde{R}_1, \tilde{R}_2, \tilde{R}_3)$ can be achieved in the partially asynchronous three senders case, $\alpha \in (0, 1)$. Using Lemma 4 it can be assumed that (R_1, R_2, R_3) and $(\tilde{R}_1, \tilde{R}_2, \tilde{R}_3)$ lie on edges of same type. Without loss of generality it can be assumed that this common type is $S = \{1, 3\}$.

Let W' and π be the 4-senders channel and the ordering in Lemma 3 for $S = \{1, 3\}$. From the proof of Lemma 3 it can be seen that W' is the first sender split version of W and $\pi = (2, 1a, 3, 1b)$. As a consequence of Lemma 3 there exist R_{1a}, R_{2a} with $R_{1a} + R_{2a} = R_1$, $\tilde{R}_{1a}, \tilde{R}_{2a}$ with $\tilde{R}_{1a} + \tilde{R}_{2a} = \tilde{R}_1$ and input distributions $p_{X_{1a}}, p_{X_{1b}}, p_{\tilde{X}_{1a}}, p_{\tilde{X}_{1b}}$ such that $(R_{1a}, R_{1b}, R_2, R_3)$ and $(\tilde{R}_{1a}, \tilde{R}_{1b}, \tilde{R}_2, \tilde{R}_3)$ are those vertices of $D(\mathcal{R}[W'; p_{X_{1a}} \times p_{X_{1b}} \times p_{X_2} \times p_{X_3}])$ and $D(\mathcal{R}[W'; p_{\tilde{X}_{1a}} \times p_{\tilde{X}_{1b}} \times p_{\tilde{X}_2} \times p_{X_3}])$ respectively, which can be described by ordering π .

If a sender is split, then the delays of the two virtual senders are equal to the delay of the original sender. Hence it is enough to prove that $\alpha(R_{1a}, R_{1b}, R_2, R_3) + (1 - \alpha)(\tilde{R}_{1a}, \tilde{R}_{1b}, \tilde{R}_2, \tilde{R}_3)$ can be achieved for channel W' when the delay system is the following: $D_{1a}(n) = D_{1b}(n) = D_2(n)$ and $D_3(n)$ are independent and uniformly distributed on the set $\{0, 1, \dots, n - 1\}$.

Note that the coordinates of the 4-tuple $(R_{1a}, R_{1b}, R_2, R_3)$ can be described as follows: $R_2 = I(X_2 \wedge Y), R_{1a} = I(X_{1a} \wedge Y | X_2)$, $R_3 = I(X_3 \wedge Y | X_{1a}, X_2)$ and $R_{1b} = I(X_{1b} \wedge Y | X_{1a}, X_2, X_3)$, where the joint distribution of $(X_{1a}, X_{1b}, X_2, X_3, Y)$ is determined by the product input distribution $p_{X_{1a}} \times p_{X_{1b}} \times p_{X_2} \times p_{X_3}$ and the channel transition W' . Similarly the coordinates of the 4-tuple $(\tilde{R}_{1a}, \tilde{R}_{1b}, \tilde{R}_2, \tilde{R}_3)$ can be described by the equations: $\tilde{R}_2 = I(\tilde{X}_2 \wedge \tilde{Y}), \tilde{R}_{1a} = I(\tilde{X}_{1a} \wedge \tilde{Y} | \tilde{X}_2)$, $\tilde{R}_3 = I(X_3 \wedge \tilde{Y} | \tilde{X}_{1a}, \tilde{X}_2)$ and $\tilde{R}_{1b} = I(\tilde{X}_{1b} \wedge \tilde{Y} | \tilde{X}_{1a}, \tilde{X}_2, X_3)$, where the joint

distribution of $(\tilde{X}_{1a}, \tilde{X}_{1b}, \tilde{X}_2, X_3, \tilde{Y})$ is determined by the product input distribution $p_{\tilde{X}_{1a}} \times p_{\tilde{X}_{1b}} \times p_{\tilde{X}_2} \times p_{X_3}$ and the channel transition W' .

Random coding argument is used, assuming without any loss of generality that αn and $(1 - \alpha)n$ are integers. The symbols of the random code-books are independent but not identically distributed random variables. The codewords of the virtual senders 1a, 1b, and sender 2 consist of two parts. The first αn symbols have distributions $p_{X_{1a}}, p_{X_{1b}}$ and p_{X_2} respectively, while the last $(1 - \alpha)n$ symbols have distributions $p_{\tilde{X}_{1a}}, p_{\tilde{X}_{1b}}$ and $p_{\tilde{X}_2}$ respectively. The symbols of codewords of sender 3 are identically distributed according to the distribution p_{X_3} . We show that with this code-book structure it is possible to achieve the rate tuple $\alpha(R_{1a}, R_{1b}, R_2, R_3) + (1 - \alpha)(\tilde{R}_{1a}, \tilde{R}_{1b}, \tilde{R}_2, \tilde{R}_3)$, by successive decoding with ordering $(2, 1a, 3, 1b)$ for channel W' if senders 1a, 1b, 2 are synchronized but sender 3 is not synchronized with them.

Note that we do not assume that the receiver knows the delays.

First the receiver decodes the codewords of sender 2. The situation is now more complicated than in case of identically distributed symbols. From the receiver's point of view the codewords of the second sender go through two different channels according to the different symbols of the codewords of the other senders. From the fact that the senders 1a, 1b, 2 are synchronized the receiver knows that the first αn consecutive symbols of codewords go through the channel

$$W^2(y|x_2) = \sum_{x_{1a} \in \mathcal{X}_1} \sum_{x_{1b} \in \mathcal{X}_1} \sum_{x_3 \in \mathcal{X}_3} p_{X_{1a}}(x_{1a}) p_{X_{1b}}(x_{1b}) p_{X_3}(x_3) W'(y|x_{1a}, x_{1b}, x_2, x_3), \quad (64)$$

and the last $(1 - \alpha)n$ consecutive symbols of codewords go through the channel

$$\tilde{W}^2(y|x_2) = \sum_{x_{1a} \in \mathcal{X}_1} \sum_{x_{1b} \in \mathcal{X}_1} \sum_{x_3 \in \mathcal{X}_3} p_{\tilde{X}_{1a}}(x_{1a}) p_{\tilde{X}_{1b}}(x_{1b}) p_{X_3}(x_3) W'(y|x_{1a}, x_{1b}, x_2, x_3). \quad (65)$$

The decoder does the following. As in Theorem 2 the n tuples $(Y_{-n+1}, \dots, Y_0), \dots, (Y_0, \dots, Y_{n-1})$ are examined. The receiver decodes the s -th codeword as the 0-th message of sender 2 if there exists an n tuple of examined output $(Y_{-n+i}, \dots, Y_{i-1})$ such that the first αn symbols of the s -th codewords are jointly typical with the first αn symbols of $(Y_{-n+i}, \dots, Y_{i-1})$ and the same is true for the last $(1 - \alpha)n$ symbols according to channels W^2 and \tilde{W}^2 respectively, and there are no other codewords with this property. With the shifted versions of this decoding technique the receiver also decodes the $-2, -1, 1, 2$ -th messages of sender 2 to ensure the decoding of the 0-th message of sender 1b in the last successive step. Note also that implicitly the receiver learns the delay of sender 2 (See Remark 8).

In the following successive step the receiver decodes the $-2, -1, 0, 1, 2$ -th codewords of sender 1a considering typicality according to the channels

$$W^{1a}(y, x_2|x_{1a}) = \sum_{x_{1b} \in \mathcal{X}_1} \sum_{x_3 \in \mathcal{X}_3} p_{X_{1b}}(x_{1b}) p_{X_2}(x_2) p_{X_3}(x_3) W'(y|x_{1a}, x_{1b}, x_2, x_3) \quad (66)$$

and

$$\tilde{W}^{1a}(y, x_2|x_{1a}) = \sum_{x_{1b} \in \mathcal{X}_1} \sum_{x_3 \in \mathcal{X}_3} p_{\tilde{X}_{1b}}(x_{1b}) p_{\tilde{X}_2}(x_2) p_{X_3}(x_3) W'(y|x_{1a}, x_{1b}, x_2, x_3). \quad (67)$$

In the third successive step the decoder deals with sender 3. Note that sender 3 is not synchronized with senders 1a, 1b, 2, hence using the $-2, -1, 0, 1, 2$ -th codewords of the senders 2 and 1a the receiver can decode (surely) just the $-1, 0, 1$ -th codewords of sender 3. It is also true in this case that the symbols of the codewords of sender 3 go through two different channels:

$$W^3(y, x_{1a}, x_2|x_3) = \sum_{x_{1b} \in \mathcal{X}_1} p_{X_{1a}}(x_{1a}) p_{X_{1b}}(x_{1b}) p_{X_2}(x_2) W'(y|x_{1a}, x_{1b}, x_2, x_3) \quad (68)$$

and

$$\tilde{W}^3(y, x_{1a}, x_2|x_3) = \sum_{x_{1b} \in \mathcal{X}_1} p_{\tilde{X}_{1a}}(x_{1a}) p_{\tilde{X}_{1b}}(x_{1b}) p_{\tilde{X}_2}(x_2) W'(y|x_{1a}, x_{1b}, x_2, x_3). \quad (69)$$

But there is an essential difference: due to the assumption on the delays it is not known which part of the codewords goes through the channel W^3 , it can be any αn consecutive symbols of the codewords. Here the word consecutive is understood modulo n . When the receiver is looking for typicality, n^2 joint typicality examinations are necessary according to the n possible positions of the separating line of the two possible channels and the possible codeword positions.

In the final successive step the receiver decodes the 0-th codeword of sender 1b considering typicality according to the channels

$$W^{1b}(y, x_{1a}, x_2, x_3|x_{1b}) = p_{X_{1a}}(x_{1a}) p_{X_2}(x_2) p_{X_3}(x_3) W'(y|x_{1a}, x_{1b}, x_2, x_3) \quad (70)$$

and

$$\tilde{W}^{1b}(y, x_{1a}, x_2, x_3|x_{1b}) = p_{\tilde{X}_{1a}}(x_{1a}) p_{\tilde{X}_2}(x_2) p_{\tilde{X}_3}(x_3) W'(y|x_{1a}, x_{1b}, x_2, x_3). \quad (71)$$

Following the calculation method of Theorem 2 it can be calculated that the rate tuple $\alpha(R_{1a}, R_{1b}, R_2, R_3) + (1 - \alpha)(\tilde{R}_{1a}, \tilde{R}_{1b}, \tilde{R}_2, \tilde{R}_3)$ is achievable with this method. Note that a genie added version of the model is necessary to a fully rigorous error estimation, as in [7]. It is also crucial that the number of joint typicality examinations is polynomial in n . One part of the complete calculation can be found in Appendix C. ■

7 Summary

This paper provides a general converse for the asynchronous multiple access channels which depends on the distribution of the delays. Interesting examples of capacity regions between the simple union and its convex closure are given. These previously unknown capacity regions are: when the set of possible delays are restricted to the even numbers, and when 2 out of 3 senders are synchronized but the third is not. These examples show, that the theory of asynchronous systems is far from complete.

It would be interesting to define a modified model where there are pauses in the communication and hence the delays could change. Further it would be important to model practical delay systems, and analyze systems with several groups of senders where each group is synchronized. Moreover, it would be interesting to solve the open problem proposed in [6] in which the possible delays are in the set $\{0, 1, \dots, \alpha n\}$ for some $\alpha \in (0, 1)$.

8 Appendix A

The coding theorem for totally asynchronous MAC (with two senders) was first stated in [14], for the case when receiver knows the delays. The theorem was stated for maximal error but the converse actually proved even for average error. While the paper [14] is hard to read, with the help of the reviewers of the previous version of this paper we have checked that the converse proof is correct, up to a minor gap pointed out after eq. (19). The achievability proof in [14] is not addressed here since more accessible proofs have been published since then. In [10] the same capacity region was claimed to be achievable also when the receiver was uninformed. However, in the delay-detection part of the proof in Appendix 1 of [10] there is a gap, in eq. (12b) an independence is assumed that need not hold when the examined n -block of channel input symbols consists of two parts, a codeword part and a sync sequence part. The other part of the proof addresses decoding the sent codewords when the delay is already known. This part is correct, giving rise to a valid achievability proof in the case of informed receiver. Another such proof was given in [7] via rate splitting and successive decoding, for any number of senders. Achievability in the uninformed receiver case has not been revisited until recently. The error in [10] mentioned above was corrected in [6]. Our approach to delay detection differs from that of [10] and [6] not so much in the latter employing a sync sequence (the first version of [5] also did that) but rather in our using a technique from [8] to bound the probability of delay detection error, while in [6] a different clever method is used to fill the gap in [10].

9 Appendix B - Proof of Theorem 1

Let $S \subset [K]$. We will derive a bound for $R(S)$. As in Section 3, take a window of the receiver consisting of $N + 1$ n -length blocks $\mathbf{Y}^{N+1} = \{Y_0, Y_1, \dots, Y_{n(N+1)-1}\}$ and the codewords having index between 1 and N from all senders (they are fully covered by this window). Recall that \mathbf{D} denotes the delay vector and $\mathbf{X}_{B, i+D_B}$ denotes the random vector with components $X_{l, i+D_l}$, $l \in B$ where $B \subset [K]$. Denote by \mathbf{X}^{sw} the $2K$ input codewords which overlap with the beginning and end of \mathbf{Y}^{N+1} . Then

$$NnR(S) = \tag{72}$$

$$= H(\mathbf{M}_S^N) \tag{73}$$

$$= I(\mathbf{M}_S^N \wedge \hat{\mathbf{M}}_S^N) + H(\mathbf{M}_S^N | \hat{\mathbf{M}}_S^N) \tag{74}$$

$$\leq I(\mathbf{M}_S^N \wedge \hat{\mathbf{M}}_S^N) + Nn\varepsilon_n \tag{75}$$

$$\leq I(\mathbf{X}_S^N \wedge \mathbf{Y}^{N+1}, \mathbf{X}^{sw}, \mathbf{D}) + Nn\varepsilon_n \tag{76}$$

$$= H(\mathbf{X}_S^N | \mathbf{D}) - H(\mathbf{X}_S^N | \mathbf{Y}^{N+1}, \mathbf{D}) + H(\mathbf{X}_S^N | \mathbf{Y}^{N+1}, \mathbf{D}) - H(\mathbf{X}_S^N | \mathbf{X}^{sw}, \mathbf{Y}^{N+1}, \mathbf{D}) + Nn\varepsilon_n \tag{77}$$

$$\leq H(\mathbf{X}_S^N | \mathbf{X}_{S^c}^N, \mathbf{D}) - H(\mathbf{X}_S^N | \mathbf{Y}^{N+1}, \mathbf{X}_{S^c}^N, \mathbf{D}) + I(\mathbf{X}^{sw} \wedge \mathbf{X}_S^N | \mathbf{Y}^{N+1}, \mathbf{D}) + Nn\varepsilon_n \tag{78}$$

$$= I(\mathbf{X}_S^N \wedge \mathbf{Y}^{N+1} | \mathbf{X}_{S^c}^N, \mathbf{D}) + Kn \log |\mathcal{X}| + Nn\varepsilon_n \tag{79}$$

$$\begin{aligned}
&= H(\mathbf{Y}^{N+1} | \mathbf{X}_{S^c}^N, \mathbf{D}) - H(\mathbf{Y}^{N+1} | \mathbf{X}_S^N, \mathbf{X}_{S^c}^N, \mathbf{D}) + \\
&\quad + Kn \log |\mathcal{X}| + Nn\varepsilon_n
\end{aligned} \tag{80}$$

$$\begin{aligned}
&= H(\mathbf{Y}^{N+1} | \mathbf{X}_{S^c}^N, \mathbf{D}) + Kn \log |\mathcal{X}| + Nn\varepsilon_n - \\
&\quad - \sum_{j=0}^N \sum_{i=0}^{n-1} H(Y_{nj+i} | \mathbf{Y}_1^{nj+i-1}, \mathbf{X}_S^N, \mathbf{X}_{S^c}^N, \mathbf{D})
\end{aligned} \tag{81}$$

$$\begin{aligned}
&\leq H(\mathbf{Y}^{N+1} | \mathbf{X}_{S^c}^N, \mathbf{D}) - \\
&\quad - \sum_{j=1}^{N-1} \sum_{i=0}^{n-1} H(Y_{nj+i} | \mathbf{X}_{[K], nj+i+D_{[K]}}, \mathbf{D}) + \\
&\quad + Kn \log |\mathcal{X}| + Nn\varepsilon_n.
\end{aligned} \tag{82}$$

Now introduce the a random variable \tilde{Y}_i linked to the random variables $X_{1,i \oplus D_1}, X_{2,i \oplus D_2}, \dots, X_{K,i \oplus D_K}$ by the channel W for all $i \in \{0, 1, \dots, n-1\}$. Then (82) is continued as

$$\begin{aligned}
&\leq \sum_{i=0}^{n-1} \left[(N-1) H(\tilde{Y}_i | \mathbf{X}_{S^c, i \oplus D_{S^c}}, \mathbf{D}) + \sum_{i=0}^{n-1} H(Y_i) + \right. \\
&\quad + \sum_{i=Nn}^{Nn+n-1} H(Y_i) - (N-1) H(\tilde{Y}_i | \mathbf{X}_{[K], i \oplus D_{[K]}}, \mathbf{D}) \left. \right] + \\
&\quad + Kn \log |\mathcal{X}| + Nn\varepsilon_n
\end{aligned} \tag{83}$$

$$\begin{aligned}
&\leq (N-1) \sum_{i=0}^{n-1} I(\mathbf{X}_{S, i \oplus D_S} \wedge \tilde{Y}_Q | \mathbf{X}_{S^c, i \oplus D_{S^c}}, \mathbf{D}) + \\
&\quad + 2n \log |\mathcal{Y}| + Kn \log |\mathcal{X}| + Nn\varepsilon_n
\end{aligned} \tag{84}$$

Dividing by Nn and going with N to infinity give

$$R(S) \leq I(\mathbf{X}_{S, Q \oplus D_S} \wedge \tilde{Y}_Q | \mathbf{X}_{S^c, Q \oplus D_{S^c}}, Q, \mathbf{D}) + \varepsilon_n.$$

This proves Theorem 1.

10 Appendix C - Some calculations to Theorem 5

The coding/decoding task of sender 3: the random code-book of the third sender consists of i.i.d symbols with distribution p_{X_3} . This code-book contains $2^{(n\alpha I(X_3 \wedge Y | X_{1a}, X_2) + (1-\alpha)I(X_3 \wedge \tilde{Y} | \tilde{X}_{1a}, \tilde{X}_2)) - \delta'}$ codewords. αn consecutive symbols of an input codeword go through channel W^3 , while $(1-\alpha)n$ consecutive symbols of the codeword go through channel \tilde{W}^3 . Here 'consecutive' is understood modulo n . Let us denote by $T \subset \{0, \dots, n-1\}$ those indices when W^3 was used. T will be called separating pattern. Note that $|T| = \alpha n$ and T contains consecutive numbers. The separating pattern depends on the relative delay D between the synchronized senders 1a, 1b, 2 and the unsynchronized sender 3. The decoder sees an output flow (note that the symbols of senders 1a, 2 are also the part of the output). The decoder does not know where the codewords are separated, and does not know the separating pattern of the two possible channels. Hence the decoder should check the same output n tuples as the decoder of Section 2 when looking for joint typicality, but when it examines an output n tuple Y^n the decoder should check every possible separating pattern. We say that the s 'th codeword is typical in window Y^n relative to separating pattern T if parts of the codewords consisting of the T coordinates of $\mathbf{X}^n(s)$ and Y^n are jointly typical according to channel W^3 , and the same is true for the coordinates $T^c = \{0, \dots, n-1\} \setminus T$ according to channel \tilde{W}^3 . If s is the only codeword which is typical in all the examined output windows relative to all separation patterns, then the decoder's estimation is s for the 0'th message. Let us consider first the case when the examined output window is an output of the r 'th codeword and when the real separating pattern is T . Then in this window the r 'th codeword will be typical relative to T with probability exponentially close to 1 by classical arguments. First we show that no other codewords will be typical in this window. Let T' be any separation pattern (it can be T too). We will estimate the probability that the $s \neq r$ 'th codeword will be typical in this window relative to T' .

For any separation pattrer T' let $P_{X_3}^{T'}(x^n, y^n)$ be the joint distribution on $\mathcal{X}^{|T'|} \times \mathcal{Y}^{|T'|}$ induced by the $|T'|$ -th power of p_{X_3} and by the memoryless channel W^3 . Let $q_{X_3}^{T'}$ be the marginal of $P_{X_3}^{T'}$ on $\mathcal{Y}^{|T'|}$. Similarly let $\tilde{P}_{X_3}^{T'}(x^n, y^n)$ be the joint distribution on $\mathcal{X}^{|T'|} \times \mathcal{Y}^{|T'|}$ induced by the $|T'|$ -th power of p_{X_3} and by the memoryless channel \tilde{W}^3 . Let $\tilde{q}_{X_3}^{T'}$

be the marginal of $\tilde{P}_{X_3}^{T'}$ on $\mathcal{Y}^{|T'|}$. Furthermore, if \mathbf{x} is an n -length sequence, then $\mathbf{x}^{T'}$ will denote the vector of length $|T'|$ consisting of those coordinates of \mathbf{x} which are in $|T'|$.

$$\begin{aligned} & \Pr_{\text{cond}} \left\{ (X_1(s), \dots, X_n(s), Y_1, \dots, Y_n) \in S_n^\delta(T') \right\} \\ &= \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta(T')} p_{X_3}^n(\mathbf{x}^n(s)) \\ & \cdot \Pr_{\text{cond}} \{ (Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s) \} \end{aligned} \quad (85)$$

$$\begin{aligned} &= \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta(T')} p_{X_3}^n(\mathbf{x}^n(s)) \frac{q_{X_3}^{T'}(\mathbf{y}^{T'}) \tilde{q}_{X_3}^{T'^c}(\mathbf{y}^{T'^c})}{q_{X_3}^{T'}(\mathbf{y}^{T'}) \tilde{q}_{X_3}^{T'^c}(\mathbf{y}^{T'^c})} \\ & \cdot \Pr_{\text{cond}} \{ (Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s) \} \end{aligned} \quad (86)$$

$$\begin{aligned} &\leq \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta(T')} 2^{-n\alpha(I(X_3 \wedge Y | X_{1a}, X_2) - \delta)} \\ & \cdot 2^{-n(1-\alpha)(I(X_3 \wedge \tilde{Y} | \tilde{X}_{1a}, \tilde{X}_2) - \delta)} \\ & \cdot \frac{P_{X_3}^{T'}(\mathbf{x}^{T'}(s), \mathbf{y}^{T'}) \tilde{P}_{X_3}^{T'^c}(\mathbf{x}^{T'^c}(s), \mathbf{y}^{T'^c})}{q_{X_3}^{T'}(\mathbf{y}^{T'}) \tilde{q}_{X_3}^{T'^c}(\mathbf{y}^{T'^c})} \\ & \cdot \Pr_{\text{cond}} \{ (Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s) \} \end{aligned} \quad (87)$$

$$\begin{aligned} &\leq 2^{-n\alpha(I(X_3 \wedge Y | X_{1a}, X_2) - \delta)} \\ & \cdot 2^{-n(1-\alpha)(I(X_3 \wedge \tilde{Y} | \tilde{X}_{1a}, \tilde{X}_2) - \delta)} \\ & \cdot \sum_{(\mathbf{x}^n(s), \mathbf{y}^n) \in S_n^\delta(T')} P_{X_3}^{T'}(\mathbf{x}^{T'}(s) | \mathbf{y}^{T'}) \tilde{P}_{X_3}^{T'^c}(\mathbf{x}^{T'^c}(s) | \mathbf{y}^{T'^c}) \\ & \cdot \Pr_{\text{cond}} \{ (Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s) \} \end{aligned} \quad (88)$$

Note that $(X_1(s), \dots, X_n(s))$ is independent from the output window ($s \neq r$), hence 1 is an upper bound of the inner sum. Note also that if one should expand $\Pr_{\text{cond}} \{ (Y_1, \dots, Y_n) = \mathbf{y}^n | (X_1(s), \dots, X_n(s)) = \mathbf{x}^n(s) \}$ then the real separation pattern T should be considered. If the examined output window is related to two codewords (the r 'th and the l 'th codewords, where $r \neq l$) then the same argument works if $s \neq l$ and $s \neq r$. If one of them is equal to s then Gray's summing technique works, the derivations can be done as in Section 4a.

Acknowledgment

The preparation of this article would not have been possible without the support of Dr. Imre Csiszár. We would like to thank him for his help and advice within this subject area. The authors would like to thank the reviewers for their very useful comments.

References

- [1] R. Ahlswede, *Multi-way communication channels*. Proceedings of 2nd International Symposium on Information Theory, Tsahkadsor, Armenian SSR, 1971. Akadémiai Kiadó, Budapest, pp. 23-52.
- [2] M. Bierbaum and H. M. Wallmeier, *A note on the capacity region of the multi-access channel*, IEEE Trans. Inform. Theory, vol. IT-25, p. 484, July 1979
- [3] T. M. Cover, R. J. McEliece, E. C. Posner, *Asynchronous Multiple-Access Channel Capacity*, IEEE Trans. Inform. Theory, vol. IT-27, NO. 4, July 1981.
- [4] I. Csiszár, J. Körner, *Information theory, Coding theorems for Discrete Memoryless Systems, 2nd edition*, Cambridge University Press, 2011
- [5] L. Farkas, T. Kóli, *Capacity Region of Discrete Asynchronous Multiple Access Channels*, IEEE International Symposium of Information Theory, Aug. 2011.

- [6] Abbas El Gamal, Young-Han Kim *Network Information Theory*, Cambridge University Press, 2012
- [7] A. J. Grant, B. Rimoldi, R. L. Urbanke, P. A. Whiting, *Rate-Splitting Multiple Access For Discrete Memoryless Channels*, IEEE Trans. Inform. Theory, vol. 47, NO. 3, Mar. 2001.
- [8] R. M. Gray, *Sliding-Block Joint Source/Noisy-Channel Coding Theorems*, IEEE Trans. Inform. Theory, vol. IT-22, NO. 6, Nov. 1976.
- [9] S. Hanly, P. Whiting, *Constraints on capacity in a multi-user channel*, in Proc. 1994 IEEE Int. Symp. Information Theory, Trondheim, Norway, June 27-July 1, 1994, p.54.
- [10] J. Y. N. Hui, P. A. Humblet, *The Capacity Region of the Totally Asynchronous Multiple-Access Channel*, IEEE Trans. Information Theory, vol. IT-31, NO. 2, Mar. 1985.
- [11] H. Liao, "Multiple Access Channels", Ph.D. dissertation, Dept. Elec. Eng., Univ. Hawaii, Honolulu, 1972
- [12] B. Rimoldi *Generalized Time Sharing: A Low-Complexity Capacity-Achieving Multiple-Access Technique*, IEEE Trans. Inform. Theory, vol. 47, NO. 6, Sept. 2001.
- [13] D. Tse, S. Hanly, *Multi-access fading channels - Part I: Polymatroid structure, optimal resource allocation and throughput capacities*, IEEE Trans. Inform. Theory, vol. 44, pp. 2796-2815, Nov. 1998.
- [14] G. Sh. Poltyrev, *Coding in an Asynchronous Multiple-Access Channel*, Probl. Peredachi Inf., vol. 19, no. 3, pp. 12-21, 1983.
- [15] S. Verdú *Multiple-Access Channels with Memory with and without Frame Synchronism*, IEEE Trans. Inform. Theory, vol. 35, NO. 3, May. 1989.